UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura
www.uniud.it

PhD course in Biomedical Science and Biotechnologies

*University of Udine*

XXX cycle

# "Unusual nucleic acid structures in the RAS genes and design of anti-cancer strategies"

**PhD student** Giulia Miglietta

**Supervisor** Prof. Luigi E. Xodo

**Co-supervisor** Dott.ssa Susanna Cogoi

2018

# 1. ABSTRACT


**Background**

The RAS genes are the most mutated genes in human cancer. More than 90 % of pancreatic tumours bear mutations in the Kirsten (*KRAS*) gene, while about 30 % of bladder tumours are characterized by mutations in Harvey (*HRAS*) gene. These oncogenes encode for highly homologues hyperactivated GTPases that constitutively stimulate survival and proliferation pathways in cancer cells. Recent data have revealed that both *KRAS* and *HRAS* are regulated by a guanine-rich sequence, located immediately upstream of the transcription start site. A main feature of this regulatory sequence is its capacity to fold into a G-quadruplex structure (or G4). Research over the last 10 years provided evidence that G-quadruplex in gene promoter plays an important role in the mechanism controlling gene expression. Although many efforts have been made, the exact function of G4 DNA in gene promoter is still unclear. A recent G4-ChIP-sequenceing analysis conducted on human chromatin found that under *in vivo* conditions the *KRAS* promoter is indeed folded into a G4. Combining this discovery with recent data from our laboratory support the view that, most likely, the main function of G4 is to recruit the transcription factors for activating transcription. However, it is important to consider that the formation of G4 by a G-rich sequence leaves the complementary pyrimidine-rich strand unpaired and exposed to cellular nucleases. The risk of DNA damage by nucleases is reduced if the pyrimidine-rich strand too folds into an unusual structure called the *i*-motif (*i*M). To address this issue we have analyzed the capacity of the C-rich motif of *HRAS* to assume the *i*-motif structure. We also found that the *i*-motif is recognized by hnRNP A1, a nuclear protein that interacts also with G4 structure formed by the complementary G-rich strand (Results, Section A). Considering that pancreatic and bladder cancer cells are addicted to *HRAS* and *KRAS*, respectively, both G4 and *i*-motif structures are viewed as attractive targets for anticancer molecules, as documented by several studies. Most of the strategies so far tested, directed against the RAS protein or downstream pathways, have not given satisfactory results, suggesting that after > 20 years of research, the RAS proteins are probably "undruggable". In the light of this belief, we focused our efforts on the genes in order to design new strategies to suppress *KRAS* and *HRAS* in cancer cells. In my PhD work I pursued three strategies: (i) the use of G4-decoy oligonucleotides, opportunely engineered, to downregulate *HRAS* in bladder cancer cells (Results, Section B); (ii) the use of small molecules targeting the 5'-UTR sequence of *KRAS* mRNA which forms several RNA G-quadruplex (RG4) structures, in order to suppress *KRAS* expression in pancreatic cancer cells (Results, Section C); (iii) the use of a modified single-stranded miRNA, which is downregulated in pancreatic cancer cells, to suppress the expression of *KRAS* in pancreatic cancer cells (Results, Section D).

**Aims**

During my research work in the laboratory of biochemistry at DAME, I have first addressed the question whether the C-rich sequence close to the transcription start site of *HRAS* folds into an *i*-motif. We found that the HRAS sequence forms an *i*-motif which, *in vitro*, is stable under near physiological conditions. We also found that hnRNP A1 is able to bind tightly to the *i*-motif and unfold this unusual structure. The suppression of hnRNP A1 by shRNA, resulting in the

downregulation of HRAS, revealed a key role played by this nuclear protein in the activation of *HRAS*.

Second, we applied a decoy strategy to inhibit *HRAS* in bladder cancer cells by using modified oligonucleotides mimicking either the *i*-motif or the critical G4 structure formed by the complementary G-rich strand.

Third, to suppress oncogenic *KRAS* in pancreatic cancer cells, we followed two different approaches aiming at blocking translation: one was based on the use of small molecules that accumulate in the cytoplasm and bind to the G-quadruplexes formed in the 5'-UTR of *KRAS* mRNA, which is very rich in guanines; the other is instead based on the use of chemically modified single-stranded miRNAs, which are downregulated in pancreatic cancer, and recognize a target sequence in the 3'-UTR of KRAS mRNA.

**Results**

The results obtained in the three-years of research has produced four scientific papers, of which three have been published in Scientific Reports, ACS Medicinal Chemistry Letters and Journal of Medicinal Chemistry and the last one is in preparation. A detailed description of the results obtained are reported in the papers enclosed in the Results section. In breaf, the main aspects of each paper are here summarized: (i) "GC-elements controlling *HRAS* transcription form *i*-motif structures unfolded by heterogeneous ribonucleoprotein particle A1" (Scientific Reports, 2015). This paper demonstrates that the two contiguous C-rich strands of *HRAS* fold into *i*-motif conformations characterized under crowding conditions (PEG-300, 40% w/v) by semi-transitions at pH 6.3 and 6.7, respectively. Chromatin immunoprecipitation shows that hnRNP A1 is associated under *in vivo* conditions to the GC-elements, while EMSA demonstrated that hnRNP A1 binds tightly to the *i*-motifs. FRET and CD experiments showed that hnRNP A1 unfolds the *i*M structures. Furthermore, when hnRNP A1 is knocked out in T24 bladder cancer cells by a specific shRNA, the *HRAS* transcript level drops to 44 ± 5% of the control, suggesting that hnRNP A1 is necessary for gene activation. The sequestration by decoy oligonucleotides of the proteins (hnRNP A1 and others) binding to the *HRAS* *i*Ms causes a significant inhibition of *HRAS* transcription. All these outcomes suggest that *HRAS* is regulated by a G-quadruplex/*i*-motif switch interacting with proteins that recognize non B-DNA conformations; (ii) "Nucleic Acid Targeted Therapy: G4 Oligonucleotides Downregulate *HRAS* in Bladder Cancer Cells through a Decoy Mechanism." (ACS Medicinal Chemistry Letters, 2015). Starting from the observation that two neighboring G-quadruplexes, located immediately upstream of the major transcription start site of *HRAS*, bind MAZ, a nuclear factor that activates transcription, we designed G4 oligonucleotides with anthraquinone insertions and locked nucleic acids (LNA) modifications mimicking one of the G-quadruplex. Luciferase, qRT-PCR, and Western blot data demonstrate that these constructs efficiently down regulate *HRAS* in T24 bladder cancer cells. The inhibitory efficiency of the G4 oligonucleotides correlates with their nuclease resistance in the cell environment. By chromatin immunoprecipitation we show that the association of MAZ to the *HRAS* promoter is strongly attenuated by the designed G4 oligonucleotides, thus suggesting that these constructs behave with a decoy mechanism; (iii) "RNA G-quadruplexes in Kirsten ras (KRAS) oncogene as targets for small molecules inhibiting translation" (Journal of Medicinal Chemistry, 2017). We found that the human *KRAS* transcript contains a G-rich 5'-UTR sequence (77 % GC)

harbouring several G4 motifs capable to form stable RNA G-quadruplex (RG4) structures that can serve as targets for small molecules. A biotin-streptavidin pull-down assay showed that 4,11-bis(2-aminoethylamino)anthra[2,3-*b*]furan-5,10-dione (2a) binds to RG4s in the *KRAS* transcript under low-abundance cellular conditions. Dual-luciferase assays demonstrated that 2a and its analogue 4,11-bis(2-aminoethylamino)anthra[2,3-*b*]thiophene-5,10-dione (2b) repress translation in a dose-dependent manner. The effect of the G4-ligands on Panc-1 cancer cells was also examined. Both 2a and 2b efficiently penetrate the cells suppressing protein p21KRAS to < 10 % of the control. The *KRAS* down-regulation induced apoptosis together with a dramatic reduction of cell growth and colony formation. In summary, we reported a strategy to suppress the *KRAS* oncogene in pancreatic cancer cells by means of small molecules binding to RG4s in the 5'-UTR of mRNA; (iv) "Potent anti Kirsten ras (*KRAS*) single-stranded miRNA oligonucleotide mimics as therapeutic agents" (in preparation). Datasets reporting microRNA expression profiles in normal and cancer cells show that miRNA 216b is aberrantly downregulated in pancreatic ductal adenocarcinoma (PDAC). Oncogenic *KRAS*, which drives the pathogenesis of PDAC, is a target of miRNA 216b. We designed single-stranded miRNA 216b mimics, engineered with unlocked nucleic acid (UNA) modifications, and found that they strongly suppress *KRAS* in PDAC cells. We also report a new delivery strategy for miRNA 216b, based on the use of palmitoyl-oleyl-phosphatidylcholine (POPC) liposomes, functionalized with a lipid-modified cell penetrating peptide (TAT) and miR-216b conjugated to a palmitoyl chain.

**Impact of the research on Cancer**

Despite the enormous advances registered in oncology, pancreatic cancer is still a highly lethal disease that poorly responds to conventional chemotherapies. New anticancer therapies are urgently needed. In this PhD thesis, we have analyzed some new possible strategies designed to inhibit *HRAS* and *KRAS* in cancer cells.

# 2. INTRODUCTION

## 2.1 RAS GENES AND CANCER

### RAS discovery

The discovery of the ras gene has provided the basis for achieving important scientific progresses in cancer biology. It has opened new horizons and given important answers in the challenging struggle against cancer, the disease of modern society.
The potent oncogenic transformation induced by the Harvey and Kirsten murine sarcoma viruses has been detected for the first time in 1964 and 1967, respectively [1, 2]. These innovative studies provided a first glimpse of oncogenic genetic elements, which have been successively called and identified as *HRAS* and *KRAS*. From late 1970, Scolnick and collaborators described the cellular origin of the viral *HRAS* and *KRAS* genes [3, 4]. They demonstrated that the *ras* genes encode for 21 kDa proteins [5], binding to GDP and GTP [6], which are associated to the plasma membrane [7]. The finding that certain human cancers are not caused by infection of oncogenic virus [8] directed the research in the identification of the processes involved on cancer development. In 1979, Weinberg and his group showed that the DNA isolated from chemically transformed rodent fibroblasts was responsible of the morphological transformation of NIH/3T3 mouse fibroblasts [9]: a finding supported by several other studies [10-14]. Scientists have been spending energies for years trying to identify and isolate genes responsible of transformation in NIH/3T3 cells. In 1982, two independent studies showed that the transforming genes were exactly those identified years before as Kirsten and Harvey viruses [15, 16]. Strikingly, at the end of the same year, it has been illustrated how a single missense mutation in codon 12 drives HRAS oncogenic activation in EJ/T24 bladder carcinoma cell line. Identical point mutations were found in the viral *HRAS* and *KRAS* genes [17, 18]. At the same time, another gene family member was discovered in neuroblastoma derived DNA and called *NRAS* [19, 20]. The remarkable role of *ras* genes in cancer has been lately validated through the identification of mutations in these genes in patient tumors. The discovery that one single mutation is able to drive the process of tumorigenesis in several cancers, have stimulated a continuous research overt last 30 years on the *ras* genes.

### RAS family and RAS isoforms

The RAS proteins are low molecular weight guanosine triphosphatases (GTPases). There are more than 150 small GTPases, classified according to their primary structure into 5 distinct subfamilies: RAS, RHO/RAC, RAB, ARF and RAN. The GTPases of the RAS subfamily share a conserved G–box-GTP/GDP-binding motif and the common ability to hydrolyze GTP into GDP. The proteins function as a molecular switch and oscillate between an activated GTP-bound state and an inactivated GDP-bound state. The conformational states depend on structural changes occurring in the two motile regions: switch I and switch II [21]. These regions are involved in the interaction with a variety of regulatory proteins. RAS proteins regulate a broad range of signalling pathways that transport into the cells extracellular stimuli. The RAS family is composed by three different isoforms: Harvey ras (*HRAS*), Kirsten ras (*KRAS*) and Neuroblastoma ras (*NRAS*). They encode proteins with a high homology in terms of amino acid (80%): the major difference is in the C-

terminal region, defined as hypervariable region (HVR) [22] (**Figure 1**). In the past, their high homology, ability of the isoforms to transform the cells and the fact that they activate the same downstream cellular effectors, led to the wrong assumption that the RAS isoforms have an interchangeable role *in vivo*. Several studies revealed that the expression of the RAS isoforms is almost ubiquitous, but the extent of expression is tissue specific [23-25]. Moreover, they are differently expressed during developmental stages [26, 27]. Mammalian *RAS* genes have different expression patterns during prenatal and postnatal development. *HRAS* is highly expressed in skin and skeletal muscles, *KRAS* in colon and thymus and *NRAS* in male germinal tissue [28]. To carry their biological functions RAS proteins are associated to the cell membrane. [29, 30]. The pathway activating RAS is mediated by tyrosine kinases receptors (RTKs) bound to the plasma membrane [31]. RAS cytosolic and inactive precursors need a series of post-translational modifications in order to acquire biological activity. Lipid modifications to the CAAX motif in the HVR (**Figure 1**) by the farnesyltransferase (FTase) are responsible for RAS localization [32, 33]. The therminal X of the CAAX motif represent methionine, serine, leucine or glutamine and it has a role in determing lipidic modification (farnesylation or geranylation) [34]. Post-translational modifications in the HVR explain RAS isoforms distinct localization and indeed different pattern of engagement [33]. For example, it has been found that FTase inhibition results with prenylation of *KRAS* and *NRAS* while not for *HRAS* [35]. Knockout studies in mice have been useful for understanding the different roles of the RAS genes. The abrogation of *KRAS* is lethal during mice embryogenesis, underling its pivotal role in development [36, 37]. Differently, when in the *KRAS* locus, *KRAS* was replaced with *HRAS*, the mice developed cardiac abnormalities in adult age [38]. These observations suggested different roles of the RAS isoforms: for instance, *HRAS* is not able to replace completely *the KRAS* function. *NRAS* activity is essential during mouse development, growth and fertility [39]. Nevertheless, both *NRAS* and *KRAS* are required for SV40T Ag-induced transformation in mouse embryonic fibroblasts showing different functions related to a different signalling pathways engagement. Wild type *NRAS* is involved in cell adhesion, through RAF and RhoA activation. On the other side, *KRAS* plays a role in coordinating cell motility, engaging AKT and Cdc42 [40].

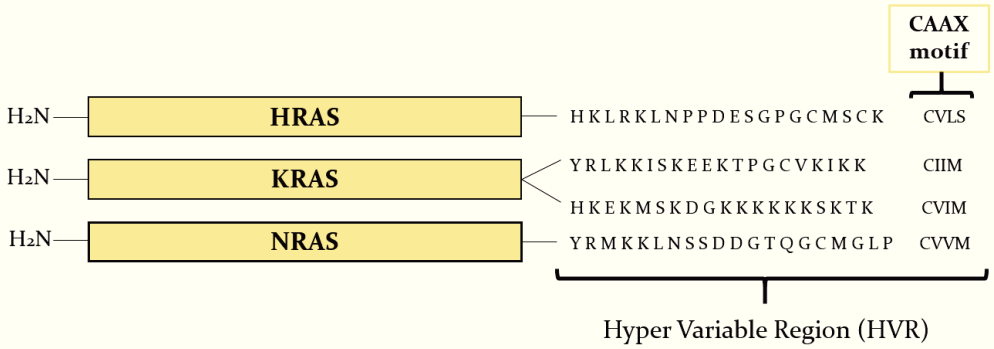



**Figure 1.** RAS protein primary sequences.

These outfindings disclosed the important role of RAS localization in defining RAS effectors engagement and specific signalling pathway activation.

RAS proteins control several pathways involved in cellular processes such as proliferation, differentiation and cell death. They are the crossroads of an enormous number of cellular signaling

network and despite different tissue-specific expression and cellular localization RAS protein share a multitude of upstream activators and downstream effectors.

**RAS upstream regulators**

The RAS signalling pathway is an elaborated network activated by a wide range of inputs. It is responsible of the transduction of extracellular signals through the activation of membrane receptors (tyrosine kinases, integrins, heterotrimeric G-proteins and cytokine receptors) [41]. Growth factors like insulin, epidermal growth factor (EGF) [42] and nerve growth factor (NGF) [43] promote the binding of GTP to the RAS protein. These factors promote conformational changes on the protein forming docking sites for adaptor molecules such as the guanine nucleotide-exchange factors (GEFs) [44]. The activation of protein RAS has profound effects on the cell. For example, activated protein RAS promotes via the EGF signaling proliferation in many types of cancer and non cancer cells. Differently, NGF-dependent activation of RAS signaling promotes cell differentiation [43, 45]. Upon receptor activation, the major RAS regulators are GEFs and GAPs (GTPases activating protein). GEFs activation promotes the dissociation of guanine nucleotides from protein RAS. RAS-GEFs are categorized into three main classes: SOS, RAS-GRF and RAS-GRP. All these proteins are characterized by the CDC-25 catalytic domain and by the N-terminal RAS exchange factors domain. Well studied GEF proteins are SOS1 and SOS2. SOS family's members display a DB1 homologous domain (DH) fundamental in managing the guanine nucleotide exchange [46]. Adaptor proteins present in the cytoplasm such as GRB2 recognize and recruit SOS to the RAS protein bound to the cell membrane [47], thus facilitating nucleotide exchange and activating RAS.
The second class of RAS regulators are the GAP proteins, which are responsible of a rapid RAS inactivation. These proteins strongly accelerate the rate of RAS-GTP hydrolysis [48]. Several GAPs have been identified: $p120^{GAP}$, NF1, GAP m, GAP III and CAPRI [46-49]. $P120^{GAP}$ has been identified as the first RAS-GAP and it displays a catalytic domain binding to the RAS effector domain. The N-terminal region is involved in modulating the catalytic activity and interacting with putative RAS downstream effectors. The tumor suppressor gene neurofibromatosis type I (NFl), that shows a sequence homology to $p120^{GAP}$ , encodes for a protein that inhibits RAS through GTP hydrolysis [50,51]. Mutation in NF1 are associated to an increased risk of cancer [52]. The fundamental role of this class of protein is also supported by the fact that RAS oncogenic mutations often abolish the interaction to GAPs causing a RAS hyperactivation responsible for tumorigenesis and cancer hallmarks.

**RAS downstream effector pathway**

The RAS downstream signal is propagated mainly through three different signalling cascades: 1) RAF/MEK/ERK (MAPK pathway) stimulating cell proliferation, 2) PI3K/AKT cascade regulating cell survival and 3) RAL guanine nucleotide exchange factor (RAL-GEFs)/RAL family GTPases signalling. RAF kinase is the first discovered RAS effector (1993) followed by PI3K and RAL-GEFs. In the years, RAS effector number has been grown and currently, apart from RAF, PI3K, RAL-GEFs and $p120^{GAP}$, it includes: RIN1, TIAM, AF6, Nore1, PLCε and PKCξ. The RAS/MEK/ERK cascade is the best described and characterized RAS effector signaling pathway [53]. The three serine/threonine kinases, A-RAF, B-RAF and RAF-1, phosphorylate and activate MEK1 /MEK2 that,

in turn, activate ERk1/ERK2 [54]. ERK kinases can activate both cytosolic and nuclear substrates, including JUN and ELK1 transcription factors that are involved in FOS expression. JUN and FOS, through AP-1 activation, are able to control cyclin D and thus cell-cycle progression [55]. The second important RAS downstream effector is PI3K. The PI3K activated state is composed by regulatory p85 and catalytic p110 subunits. Among the RAS isoforms, HRAS displays a capacity to activate PI3K higher than KRAS [56]. Once activated, PI3K, through PIP2 phosphorylation, generates phosphatidylinositol (3, 4, 5) triphosphate (PIP3) that recruits PDK1 to plasma membrane and activates AKT [57, 58]. The three different AKT isoforms (AKT1, AKT2 and AKT3) are associated to different biological processes. AKT1 is responsible of cell proliferation and survival, AKT2 promotes insulin mediated metabolic responses, and AKT3 deficient mice displayed diminished cell size and numbers [58]. PI3K also activates RAC, a RHO family GTPase that plays an important role in RAS-oncogenic transformation [59]. Lastly, RAS can also interact and activate RAL-GEFs family necessary for RAL GTPases activity. RAL-GEFs family is composed by 4 members: RAL-GDS, RGL, RGL2/RLF and RGL3 [60]. This class of protein share a high homology in three different domains: CDC25 domain, that *in vitro* has been shown to be sufficient for the catalytic activity [60], RAS exchange motif (REM) and RAS binding domain (RBD)[61]. RAS and RAL-GEFs interaction permits the complex localization to the plasma membrane and it drives RAL GTPase activation. Two RAL GTPase genes are RAL-A and RAL-B that share 80% of homology and are ubiquitously expressed in humans. Evidences showed that RAL-A is required for the proliferation of transformed cells, while RAL-B is needed for cell survival in transformed cell, but not in normal ones [62]. RAL proteins are able to interacts with a large numbers of downstream effectors (as Sec5, Filamin, RAL-BP1 or ZONAB) and their signal is implicated in regulation of endo- or exocytosis, cytoskeletal organization, cell migration and gene expression [63].

**Oncogenic RAS mutations**

RAS somatic mutations are present in about 30 % of all human cancer [64]. RAS mutations are in the tumorigenesis of several deadly cancers. In advanced stages of cancer, RAS mutations are associated to bad prognosis and more aggressive phenotype. There is a correlation between tumor types and ras mutations [65]. For instance, *KRAS* mutations are prevalent in pancreatic and colorectal cancers, while *NRAS* and *HRAS* mutations are associated to melanoma and bladder tumors [66]. The main hotspots for RAS mutations are positions 12, 13 and 61 in exon 1. These mutations impaired the correct GTP hydrolysis causing an aberrant hyperactivated protein. GTP-GDP exchange is not inhibited because RAS proteins became resistant to GAPs activity. Crystal structure of $p120^{GAP}$ catalytic domain (Gap Related Domain, GRD) and *HRAS* revealed that the highly conserved GAP-arginine-finger interacts to the RAS phosphate binding loop (P loop) [67]. It has been found that replacing Gln 61 with any amino acids, except Glu, results in the inhibition of GTP hydrolysis [67]. Yet, when Gly 12 is replaced with another amino acids (proline excluded), the affinity of GAPs or protein RAS decreases significantly. The phenotype of the transformed cells depends on the type of mutation [68]. Amino acids substitutions in positions 12 and 13 are not favourable for the physiological RAS GTP hydrolysis, because of a steric block that prevents GAP's arginine finger to enter into the GTPase site [67, 69, 70]. The mutated and hyperactivated RAS proteins produce dysregulated downstream signallings involved in cell growth, survival and differentiation.

***HRAS oncogene in bladder cancer***

Bladder cancer is the fifth most frequent malignancy in Europe and USA and the second most frequently diagnosed genitourinary tumor after prostate cancer [71]. It is the seventh reason of death among cancer in men and the eighth in women [72]. Clinically, 75% to 85% of bladder tumors have an early-stage diagnosis, explaining the lower level of mortality respect other tumors [73]. The identification of HRAS as a human oncogene has been done in T24 bladder carcinoma [74], anyway *HRAS* mutations in bladder carcinomas are not higher than 30 % [75, 76]. G12V mutation is the predominant substitution (60 % of total mutations) followed by G12D (8 %) and Q61R (7 %) [77]. Viola and collaborators demonstrated that *HRAS* levels are increased in premalignant cells (dysplastic lesions) and in high grade carcinomas [78]. Moreover, pathological significance of *HRAS* mRNA overexpression in bladder cancer have been proven [79]; *HRAS* overexpression seems to be involved in early cancer development, but not in tumor progression [75]. In these years several anti-*HRAS* therapies have been developed. For example, anti-*HRAS* ribozyme has been used in cell lines and mouse cancer model. They demonstrated antitumor efficacy through malignant phenotypes reversion and tumor growth inhibition both *in vitro* and *in vivo* [79]. Alternative strategies based on DNA non-canonical structures within *HRAS* promoter have been developed. The fact that a critical G-rich sequence in the *HRAS* promoter forms a G-quadruplex with regulatory functions, led to develop of a decoy strategy: oligonucleotides mimicking the promoter G4 of HRAS that sequester the transcription factors and arrest transcription [80, 81].

***KRAS oncogene in pancreatic adenocarcinoma (PDAC)***

Pancreatic adenocarcinoma (PDAC) is one of the most predominant lethal cancer; 90 % of patients with a pancreatic tumor diagnosis die for the disease [82]. In USA, 48'000 new cases have been estimate every year and 40'000 are associated with a deadly prognosis. It is the eighth and ninth cause of death for cancer in men and women respectively [83]. Surgical resection is the only potentially curative solution for this pathology, but only 15-20 % of patients are candidates for surgical operations [84]. Oncogenic *KRAS* have been associated with pancreatic cancer 28 years ago [85, 86]. The malignant transformation is a stepwise process that in 90 % of cases is driven by hyperactivating mutations in the *KRAS* gene [87]. The most common mutations occur in codon 12, 13 and 61. *KRAS* mutations play a central role in the initiation, progression and maintenance of PDAC [88-90]. *KRAS* mutations are frequently found in precursor lesions, as intraepithelial neoplasia (PanIN), supporting their pivotal role in PanIN formation [86]. Recent studies have demonstrated that *KRAS* G12D plays a critical role in the initial phase of tumor development and in reprogramming the metabolism in the transformed cells [91]. Tumor progression however requires additional tumor suppressors inactivations [92, 93]. The absence of effective *KRAS* inhibitors made difficult to investigate the dependency of tumor progression by KRAS. But, the introduction of RNA interference-based approaches allowed to obtain important findings on this issue [94]. In addition to ras, also its downstream effectors may contribute to pancreatic cancers development in mice. For example, AKT is involved in ductal structures expansion, but differently from *KRAS*, it does not drive PanINs or PDAC progression [95]. PI3K pathway is sufficient and crucial to initiate pancreatic carcinogenesis [96]. These observations strongly suggest that *KRAS* effectors differently contribute and influence malignant transformation.

**RAS cancer "addiction"**

In 2000, Bernard Weinstein proposed the concept of "oncogene addiction". He started from the observation that cancers in which cyclin D is overexpressed are particularly sensible to the depletion of cyclin D. He also observed that the knockout of cyclin D by siRNA reverted the pathological phenotype [97]. Data obtained in genetically modified mouse models, in human cancer cell lines and in clinical trials, support the idea that some tumors strongly depend on one or more genes to maintain the malignant phenotype. Although a large number of genetic aberrations are normally present in cancer, the transformed cells exhibit a marked dependency on a specific

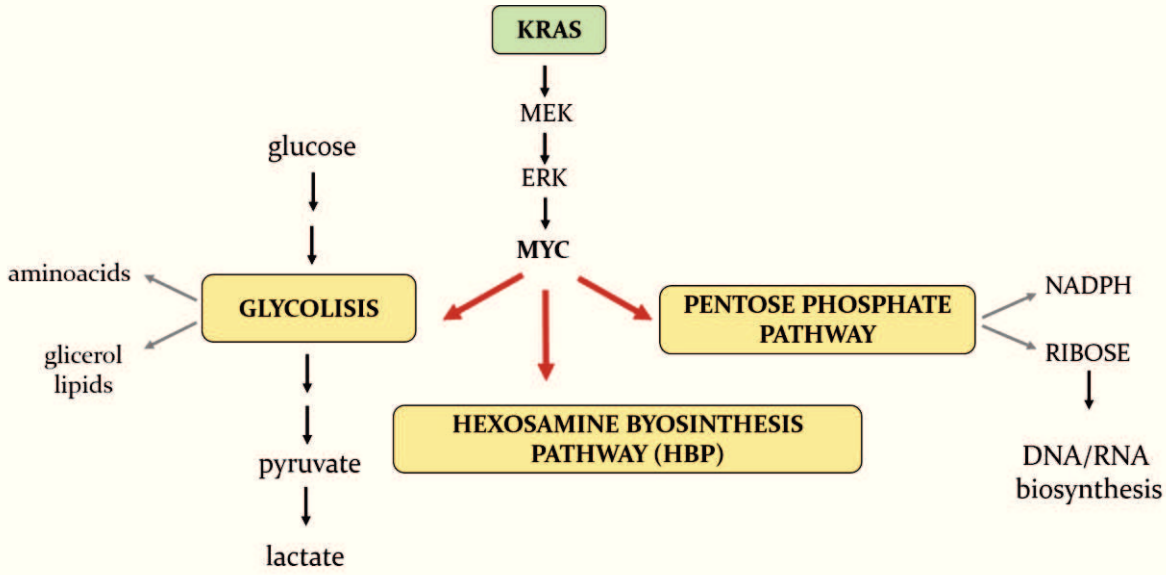



**Figure 2.** Representation of KRAS metabolism reprograming.

oncogenic signalling pathway [98, 99]. This cellular phenomenon can been seen as the "Achilles' heel" of cancer and can be exploited for the development of new anticancer drugs. Therapeutic agents that are able to reduce the expression of a critical oncogene, to which the cells are addicted, are likely to promote cell death. Tumor cell addiction has been found for several oncogenes, such as *c-MYC* in lymphomas [100, 101] and myeloma [102], *BCR-ABL* in leukemia [103], *HRAS* in myeloma [103] and *KRAS* in pancreatic cancer [104, 105]. Research of the last five years has revealed why pancreatic cancer cells are addicted to oncogenic *KRAS*. One of the main function of oncogenic *KRAS* is to orchestrate a profound reprogramming of cell metabolism, in order to fuel a high proliferation rate. Oncogenic *KRAS* affects both glucose and glutamine metabolism, and it significantly boosts the biosynthesis of ribose by directing glucose towards the pentose phosphate pathway [106]. As illustrated in the **Figure 2**, in PDAC cells KRAS increases the glycolytic flux to produce the substrates for the synthesis of amino acids and lipids. KRAS directs a larger amount of glucose towards the pentose phosphate (PP) pathway to produce ribose for the biosynthesis of the nucleotides. Moreover, the PP pathway produces the reducing power necessary to activate the anabolic pathways. PDAC cells need also a high amount of glutamine (Gln) to control the redox

balance in cancer cells [91, 107]. In fact, Gln after deamination to glutamic acid (Glu) is not oxidized into $CO_2$ in the Krebs cycle, as normally does in non-cancer cells, but it is transaminated to aspartate (Asp), which in turn is transaminated again to oxaloacetate (OAA). OAA is reduced to malate and malate decarboxylated to pyruvate with the concomitant production of reducing power, under the form of NADPH, which is necessary to control the redox balance of PDAC cells. Therefore, Gln plays an essential role to avoid a dangerous level of oxidative stress in PDAC cells that could compromise their high proliferation rate. The picture that is coming out is that oncogenic KRAS is essential for PDAC cells because it adapts the metabolism to cope with a high rate of proliferation (**Figure 2**).

**RAS inhibitors**

Mutated RAS proteins are the highest cause of cancer and are, at the same time, the most "undruggable" proteins. They have been defined "too smooth", "too floppy" or otherwise "too finicky" for drugs to bind to and block" [108]. Three decades of intensive work by the scientific community and private pharmaceutical industries have not been able to identify successful pharmacological RAS inhibitors. Despite past failures, targeting RAS is still nowadays one of the biggest challenge in therapy. Several approaches have been used and developed to repress RAS activity and downstream signalling. Past and ongoing approaches can be summarized as follows: (i) inhibitors acting directly on RAS; (ii) inhibitors of RAS membrane association; (iii) inhibitors of RAS downstream effectors; (iv) inhibitors of RAS-induced metabolism/cancer hallmarks (**Figure 3**).

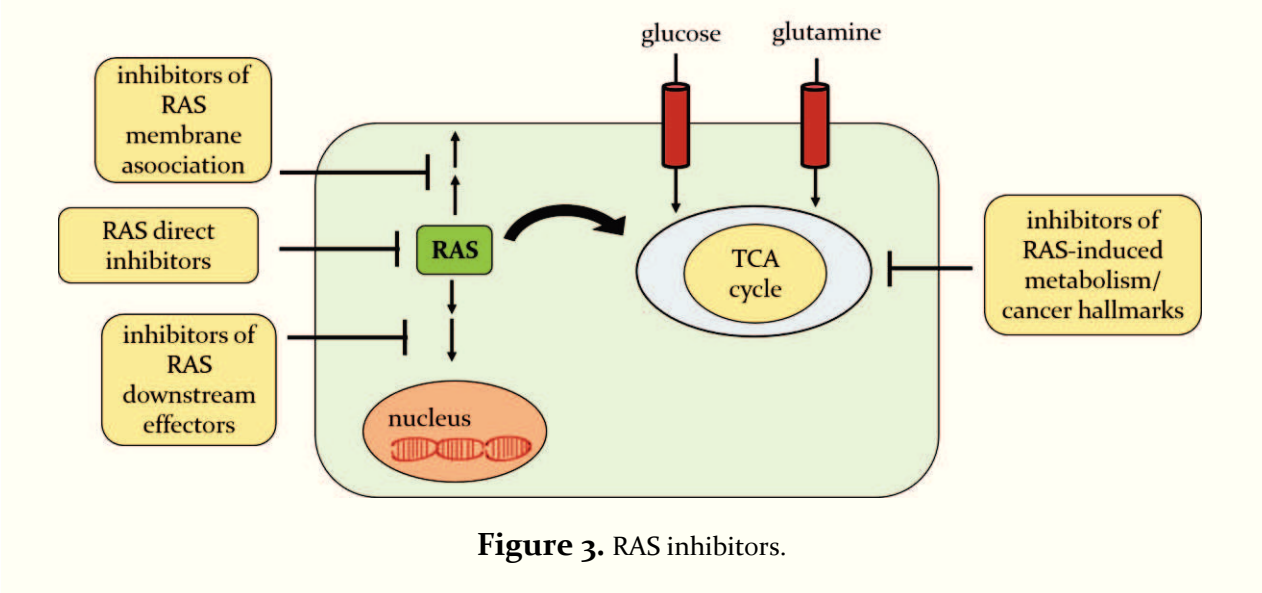



**Figure 3.** RAS inhibitors.

To compete the binding of GTP to protein RAS seems impossible because of the high affinity of protein RAS for GTP. In fact, ligands blocking the interaction between GTP and protein RAS have not yet been discovered. Neither the interaction between protein RAS and other cellular factors can be inhibited by small molecules, as the latter are unable to compete with protein-protein interaction [108]. For example, MCP1 and derivatives have been screened as inhibitors of RAS-RAF interactions. Although they seemed to inhibit RAF activation, partly reverting the malignant phenotype [109, 110] they showed to be not enough efficacious. DCAI, a compound reported by Genentech (California, USA), inhibits the GDP-GTP exchange, blocking the interaction between

RAS and SOS1 [111]. X-ray studies revealed that DCAI binds to an important pocket of protein KRAS, but unfortunately it did not exhibit a sufficient RAS-binding affinity. The structural information obtained by this and other studies were useful for further investigations [111]. In the light of these non promising approaches, the researched focused on the use of inhibitors of the farnesyl pathway binding protein RAS to the membrane. Many farnesyl transferase inhibitors (FTIs) reached clinical trials (as lonafarnib and tipifarnib) [112]. FTIs effectively reduced cell proliferation in mice and human cancer cells, but one by one, they failed when used in human clinical trials. This because *KRAS* and *NRAS*, the less known isoforms at that time, are able to activate a compensatory mechanism in the absence of farnesyl tail. They are recognized by enzymes that promote an alternative lipidic modification (prenylation) that associates protein RAS to the membrane so that it can promote its oncogenic activity. As RAS controls several downstream effectors, their targeting could be a winning approach. The strategy is promising, but important issues have to be addressed, as compensatory mechanisms can decrease the efficacy of the treatment and also cross talk between RAS pathways can be an obstacle to downstream targeting [113]. Finally, cancer cells are characterized by an altered metabolism. Undoubtedly, RAS is a key player in promoting metabolic rewiring, becoming crucial for tumor maintenance [91, 114]. On the contrary, normal cells are not addicted to KRAS, suggesting that targeting tumor metabolic pathway could be an attractive and maybe selective anti-cancer strategy. Glutaminase (GLS) inhibitors are an example of this kind of approach. They block the catalytic conversion of glutamine to glutamate altering the redox balance in pancreatic cancer cells, thus affecting their proliferation rate [115, 116].
The lack of successes in these 30 years produced a sense of frustration, but increased knowledge on RAS structure, protein domains, RAS mediated processes, and innovative tools as computational screening persuade scientific community to continue the "hunt" to pharmacological RAS inhibitors.

Channing Der, an acclaimed scientist in RAS field, compared the RAS inhibitors discovery to the climbing of Mount Everest*: "We've got to climb it, even though it's tough."*

**Innovative anti-RAS strategies**

Blocking and repressing RAS proteins activity by small molecules has thought to be a very challenging, if not an impossible task. In this scenario, there is an imperative necessity to find feasible and alternative therapeutic approaches to inhibit RAS oncogene in cancer cells. Novel strategies focused in blocking the expression of the ras gene, impairing either transcription or translation. Non-canonical DNA and RNA structures involved in the mechanism controlling ras expression are currently being explored in order to develop alternative therapeutic strategies. Several studies pointed out that gene expression is controlled by the action of specific transcription factors that recognize unusual DNA secondary structures, G-quadruplexes (or G4), formed in a critical G-rich region located upstream of TSS, that governs ras expression [117-119]. The structure of G-quadruplex DNA, its function and its importance as therapeutic target is amply described and discussed in the next chapter. Although it is not clear the exact function of G-quadruplex DNA, its presence in gene promoter is likely to have the function of recruiting specific transcription factors to activate transcription. The *HRAS* and *KRAS* G-quadruplex structures are recognized by several transcription factors including MAZ, hnRNP A1 and SP1 and their action is strongly involved in gene expression [117, 118]. The suppression of transcription factors is expected to inhibit gene expression;

therefore a decoy strategy can be a powerful strategy to inhibit ras genes in cancer cells. The rationale of the decoy strategy is based on the use of G4 oligonucleotides, mimicking the G-quadruplex structure formed in the ras promoter, that sequester the transcription factors normally binding of the *HRAS*/*KRAS* promoter with the result that transcription is blocked. The efficacy of the decoy strategies in repressing RAS gene expression has been demonstrated [80, 120, 121].

The discovery that microRNAs (miRNAs) regulate gene expression opened new therapeutic prospectives also for the ras genes. Micro RNAs are small, non-coding RNAs that act as post-transcriptional regulators. MiRNAs are 19-24 nucleotide single strand RNAs transcribed from the genome as precursors (pre-miRNAs) and then processed as mature miRNAs through Drosha and Dicer action [122, 123]. More than 50 % of the known miRNAs are involved in human tumorigenesis and metastasis by directly targeting oncogenes or tumor suppressor mRNAs targets [124-126]. Several miRNAs are aberrantly downregulated in tumours where the ras gene are mutated: let-7 [126], miRNA 96 [127], miRNA 216b [128] and miRNA 30c [129]. As they are tumor suppressors in nature, they can be used in a replenish strategy as anti ras agents [130, 131]. Chemical modifications can be used to confer nuclease resistance and enhance the oligonucleotide therapy effect [132, 133].

## 2.2 DNA AND RNA NON-CANONICAL STRUCTURES

### 2.2.1 DNA/RNA G-quadruplex

In 1910, Bang [134] noticed that a concentrated solution of guanylic acid was able to form a gel suggesting for the first time the formation of a DNA high-order structure. Through the optical features of the gel and structure investigations, he revealed that the phenomenon was due to the formation of a helix by the guanylic acid. Starting from this observation, 50 years later Gellert and coworkers [135] using X-ray diffraction, proposed a tetrameric structure for G-quadruplex DNA (G4). Research in the last 20 or more years has provided solid experimental evidence that this non-canonical structure of DNA in specific regulatory regions of the genome plays a pivotal role in the regulation of gene expression.

**G-quadruplex topology**

G-quadruplex (G4) is a non-canonical structure that can be adopted by G-rich DNA and RNA strands. The building blocks of the structure, called G-tetrads, are four guanines that trough Hoogsten bonds between the N1, N7, O6 and N2 array in a square planar conformation. Stacking of two or more G-tetrads leads to G4 folding (**Figure 4**).

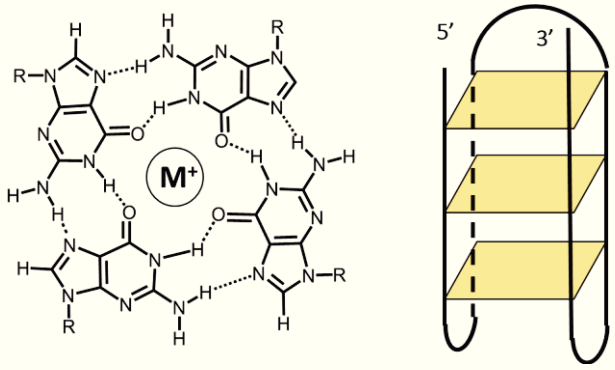



**Figure 4.** G-tetrads structure formed by four guanines with cation insertion and a representation of a G-quadruplex structure.

An important feature of the structure is that it is stabilized by the presence of free metal cations such as $K^+$, $Na^+$, $Mg^{2+}$, $Ca^{2+}$ with a strong preference for $K^+$ and $Na^+$ at typical intracellular concentrations. Biophysical and structural studies revealed that the metal ion is located in the tetrad central cavity explaining some ions preference rather than others. The cation's ray plays an important role allowing the correct insertion in the G-tetrad core where it neutralizes the electrostatic repulsions of the guanines oxygens. In general terms, G4 is defined as a stacking of at least two tetrads held together by loops formed by mixed nucleotides, including guanines present

in the sequence but not involved in tetrad formation. Anyway, the G4 structure shows a huge variety of topologies and different stability depending on: i) number of stacked G-tetrads ii) strand direction and polarity iii) loops length and location iv) sequence [136]. The G4 can be formed by one (intramolecular), two or four strands (intermolecular). In general, they are defined parallel when strands run in the same direction and all guanosines glycosidic angles are in an *anti* conformation, differently they are defined antiparallel when guanines are both in *anti* and *syn* configuration (**Figure 5**).

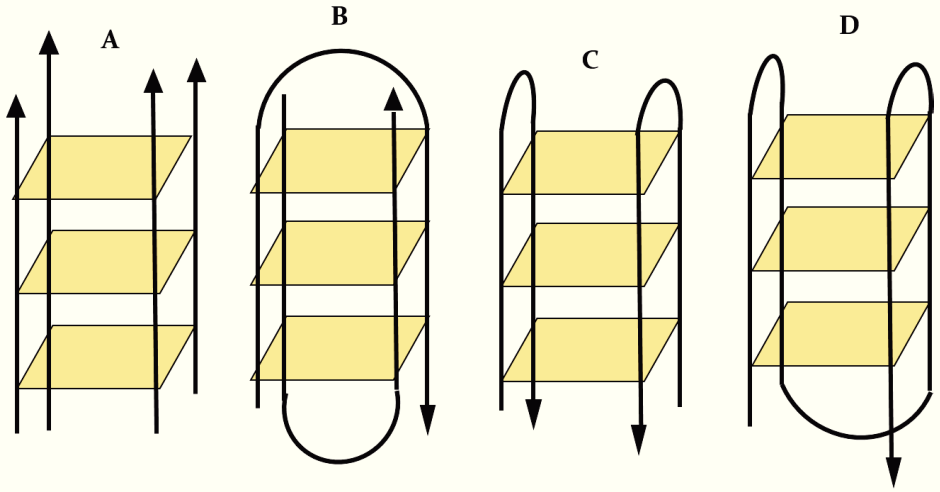


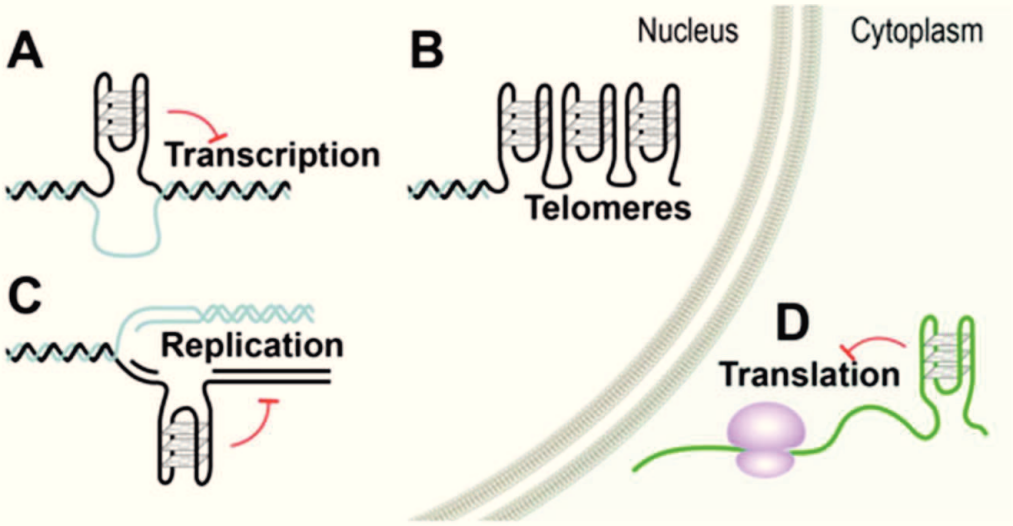

**Figure 5.** Structures and topologies of G-quadruplexes: (**A**) intermolecular with parallel strands; (**B**) bimolecular with diagonal loops; (**C**) bimolecular with lateral loops; (**D**) intramolecular with lateral loops.

**G-quadruplex biological role**

The G-quadruplex folding in synthetic sequences has been fully demonstrated by numerous biochemical and structural studies using different approaches and specific techniques. Aaron Klug, the winner of the Nobel Prize in Chemistry (1982) for the development of crystallographic electron microscopy, over 30 year ago affirmed: "If G-quadruplexes form so readily *in vitro*, Nature will have found a way of using them *in vivo*." Anyway, the existence, the detection and especially the function of the G-quadruplex in a biological context is nowadays one of the main topics of debate.

Computational predictions suggest that over 360.000 sequences can potentially assume a G4 conformation [137] and a more recent algorithm developed by Mergny and coworkers [138] suggested that the number of potential G4 is probably higher. Computational analysis lead to map G4 distribution, highlighting that G4s have not a random localization.

They tend to cluster in specific regions of the genome as promoters, telomeres, 5'-utr, near transcription factors binding sites and up- or downstream transcription start site (TSS) position. (**Figure 6**).




Reproduced from: Nucleic Acids Res. 2015 ;43 (18): 8627-8637.

**Figure 6.** G-quadruplex possible location in cells. G-quadruplex formation can occur in double stranded G-rich regions when DNA becomes transiently single stranded, during (**A**) transcription and (**C**) replication and (**B**) at the single stranded telomeric G-rich overhangs. Outside the nucleus, G-quadruplexes can also form in mRNA and (**D**) are involved in translational control. Red T-bars indicate impediments to transcription, replication and translation. Reproduced from [147].

Moreover, G4 localization and nucleotide sequences are conserved among species. The conservation is higher among mammalian species and is reduced in non-mammalian and lower species. Putative G-quadruplex-forming-sequences are present also in bacteria [139] and viruses [140]. All these observations suggest a biological function of this non-canonical conformation, emphasizing its possible role in cellular processes regulation. In 2001, Schaffitzel and colleagues provide the first evidence of G4 existence in a cellular context by using a G4-selective single chain variable fragment of antibody (scFV) [141]. They selected the scFV from a large synthetic library and through indirect immunofluorescence they showed the G4 folding within telomeres in the *Stylonychia lemnae* macronuclei. The visualization of G4s in mammalian cells through two different G4-specific antibodies 1H6 [142] and BG4 [143], developed by two independent laboratories has been an important breakthrough in the field. It is important to point out that just some months ago the scientists responsible for the generation of the 1H6 antibody reconsidered their results. They found that 1H6 cross-reacts with short single strands of thymidine (poly (T)) in denaturated DNA fibers. For this reason previous immunofluorescence interpretations using 1H6 need additional studies to affirm that 1H6 binding to DNA is due to G4 structure [144]. This fact underlines the difficulties in mapping G4 occurrence *in vivo*. Nevertheless, increasing evidences for G4 formation in both DNA and RNA are emerging and support the role of high-order structures as regulatory mechanisms in controlling biological processes such as transcription, replication, translation and telomere maintenance. A meaningful step in G4 mapping and localization *in vivo* has been the G4-dependent DNA polymerase stalling combined with the Illumina next generation

sequencing (G4-seq). This approach leads to obtain an experimentally genome-wide distribution of G4 [145]. Moreover, G4 structure-specific antibodies, used as probes for a G4-specific ChIP-seq, permit to achieve new considerations on G4 distribution and most importantly on the correlation between G4, chromatin and transcription. The immunoprecipitation of isolated chromatin through a G4-specific antibody (BG4), followed by sequencing, experimentally demonstrates the existence of 10,000 G4s in human chromatin, mainly localized in nucleosome-depleted regulatory regions. G4 number is appreciably lower than that computationally predicted or detected by G4-seq and the authors explained this discrepancy by proposing suppressive role for heterochromatin in G4 formation in human cells [146]. The G4 mapping and distribution are crucial steps to understand G4 biological role, and those findings are consistent with the idea that G4 folding has a dynamic equilibrium, which acts as a regulatory molecular switch. Genome is packaged into chromatin and DNA is normally present in the classical double stranded conformation, surely more favored than G4 folding, but nuclear G4 structures can be stabilized by superhelical stress, molecular crowding and by protein/transcriptional factors binding [147]. Additionally, specific situations, in which the classical double strand is temporary in a single strand form, could allow G4 folding; for example during DNA replication, transcription and DNA repairing. The single stranded RNA and certain telomeric G-rich single strands are other two important conditions in which non-canonical DNA/RNA folding could be preferred rather than classical Watson–Crick base paring.

### *G-quadruplex in telomeres*

Telomeres are nucleoprotein complexes present at the 3'-ends of the eukaryotic chromosomes. They are responsible for genome integrity and for chromosomes protection, especially for DNA recombination, degradation and end-to-end fusion. DNA polymerase is not able to replicate the 3'-ends because a template strand is lacking and consequently at each cell cycle telomeres are reduced by 50-100 bases in length. After reaching a critical length, the Hayflick limit [148], replication stops and cells enter in a senescence state. In immortalized and cancer cells, a compensatory mechanism that results in telomere length and integrity stabilization occurs. Telomere maintenance is due to reverse transcriptase ribonucleoprotein telomerase (hTERT) overexpression; this is the unique polymerase able to bind DNA telomeric 3' strand in order to synthetize further hexanucleotide repeats through endogenous RNA. Telomerase plays a key role in tumorigenesis because it contributes to cellular immortalization [149]. Telomerase overexpression is a cancer hallmark and it is present in 80-85% of proliferating tumor cells [150]. Human telomeric DNA consists of thousands of single-stranded tandem repeats d(TTAGGG) [151], normally protected by proteins collectively known as sheltering proteins complex. G-quadruplex folding in telomeric strands has been hypothesized and facilitated by the complementary strand absence. Four consecutive repeats can fold into stable intramolecular G4s with a huge variety of topologies and polymorphic structure proved by X-ray [152], NMR [153, 154] and other techniques [155]. In human telomeres for example two-unit repeats fold into a dimeric parallel G4 [156], whereas four-repeats form two intramolecular G4s, different only in loops arrangements; that shows how several topologies can coexist in the same sequence [157]. The first direct evidence *in vivo* of the telomeric G-quadruplex came 10 years ago studying the macronuclear telomeres of *Stylonchia*. Most importantly they showed that G-quadruplexes folding in telomeres is related to cell cycle, theorizing that G4s act as protective capping structures at chromosome ends [141]. Intramolecular G4 formed by single stranded

telomeric repeats is one of the best studied G4 due to its polymorphic structure and to its role as tumor suppressor. Telomeric G4 has been proposed as a therapeutic target in 1997. Telomerase inhibition has been achieved through G4 stabilization by 2,6-diamidoanthraquinone [158]. G4 folding, induced by ligands provides a telomerase impaired activity [159]. Telomerase inhibition scan reverts cancer cells immortalization. A wide-range of small molecules recognize and bind to telomeric G4s [160-164]. All the G4 ligands were considered as telomerase inhibitors, but for example BRACO-19, a G4 ligand discovered by Neidle and coworkers [165], shows an antitumor activity *in vivo* [166] that could be due to its ability to induce a DNA damage in telomeres [167, 168]. Tumor cells are characterized by DNA-repair deficiencies and this fact could provide ligands selectivity. The same mechanism of DNA damage induction has been proposed for an acridinium derivative [169] and for pyridostatin analogues [170].
Nowadays, telomeric G-quadruplex is a promising and valuable target for the development of novel anticancer therapies.

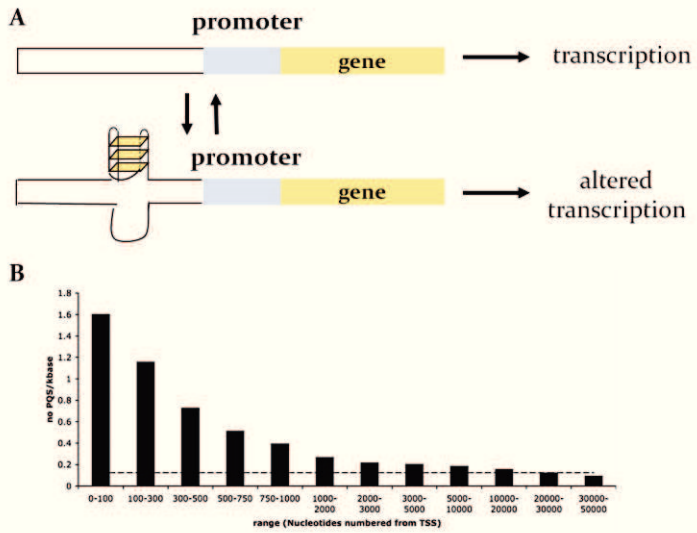
***G-quadruplex and DNA replication***

In literature a growing body of evidences demonstrates that replication can be influenced by secondary DNA structures such as hairpin [171], triplex and G-quadruplex [172, 173]. During replication, the classical double helix is resolved by helicases allowing replication fork formation required for DNA synthesis. The synthesis of the lagging strand is discontinuous and retarded and a slowed down replication can be an opportunity for a G-quadruplex formation in the retarded strand. A pre-existing G4 structure on the template strand can behave as a key regulatory factor. It can act as repressor causing a replication stall until G-quadruplex is in the folded form. On the contrary it can be a stimulatory factor acting as a protein binding site that promotes the DNA helicases recruitment. DNA helicases are molecular motors with a pivotal role in preserving genome homeostasis; they impact nucleic acids structures. Helicases are involved in human diseases, such as cancer, and ageing [174]; their deficiencies are responsible for genome instability and many of them are efficient G4 unwinders *in vitro* [175, 176]. For example Pif1 is a member of a DNA 5'-3' helicase family, which efficiently unfolds G-quadruplex substrates [177]. Pif1 in *S. cerevisiae* binds and recognizes G4 motif and it is fundamental in DNA replication. Indeed, G4 motifs slow the replication rate and, in cells where Pif1 is absent, double-strand breaks (DSBs) occur especially at G4 motifs, which are normally recognized by Pif1. G4 structures are responsible of genome instability, if they are not correctly resolved by Pif1 [178]. Another important evidence that uncorrected G4 folding/unfolding results in genomic instability, derived from the observation that cell lines with mutated FANCJ helicase are characterized by an accumulation of deletions in G-rich regions with G4 signature [179]. FANCJ function in GC-rich tracts is to resolve G-quadruplex structures to permit correct DNA synthesis; a stall in the replication process can be risky for genomic maintenance [180].

***G-quadruplex in the oncogene promoters***

Promoter regions function is to regulate transcription initiation and it has been demonstrated that they display a G4 enrichment. More than 40% of human genes exhibit putative G4 in the promoter

regions. The G4 motifs percentage is inversely correlated to TSS distance suggesting a role played by DNA non-canonical structure in transcription regulation (**Figure 7**) [181].




Adapted from: Nucleic Acids Res. 2006;35(2):406-413.

**Figure 7**. G-quadruplex in promoter regions: (**A**) Model for transcription regulation through G4 folding in promoter region; (**B**) Density of putative G4 sequence related to distance upstream TSS. Adapted from [181]

A similar enrichment of G4 structures in promoter regions has been observed also in other species as bacteria, yeast and plants [182-184]. Compelling evidences demonstrate that G-quadruplex formation is a compensatory effect induced by negative supercoiling; indeed G4 folding can be a regulating mechanism in transcription process. G4 conformation in or near promoter regions can have either a positive or a negative effect in that elaborated mechanism, one hypothesis is that it depends on which DNA strand encodes the G4 motif. If G4 is in the template strand, it can behave as a transcription repressor because the folding causes a block of transcriptional activity; on the other hand, if G4 is on the non-template strand, it could permit template maintenance as single strand promoting an effortless transcription [185]. Moreover, G4 can be useful for specific proteins recruitment. Protein unwinding activity can be responsible of transcription activation. A clear and univocal role for the G4 in transcription has not yet been proposed: several biological agents depending on circumstances and on environmental context probably influence G4 folding and functions. In literature, a large body of work shows the inhibitory effect of G4s, besides others provided data about enhanced transcription induced by G4 folding. It is not possible to rule out a general function, it is just possible to analyze each G4 in its context. It has been observed that G-quadruplexes in promoters are associated with increased transcriptional activity and enriched in oncogenes [186] arising the opportunity to consider them as powerful and promising targets in alternative anticancer therapies. The most studied G4s, on that point of view, are those present in the *c-MYC*, *KRAS* and *c-KIT* promoters [187].

***c-MYC***

Deregulation of *c-MYC* expression is usually due to gene amplifications and translocations, altered ploidy or to an aberrant transcription caused by upstream signalling abnormalities [188]. An overexpressed *c-MYC* is responsible of an oncogenic transformation both *in vitro* and *in vivo*. More than 80% of solid tumors including gastrointestinal, ovarian, breast cancer and many non-Hodgkin's lymphoma tumors are associated with *c-MYC* overexpression. Targeting *c-MYC* at the protein level is not simple, thus the demonstration, in 2002, of the G-quadruplex folding in the nuclease hypersensitive element (NHE) offered an alternative way to modulate oncogenic *c*-MYC expression. It is one of the first oncogenes whose promoter was amply and deeply investigated. G-quadruplex structure is formed by 27-nt located between -142 and -115 bp upstream the TSS in NHE III [189]. The *in vitro* structure has been resolved by NMR studies [190, 191]. *c-MYC* G-quadruplex behaves as a repressor element: its formation and its stabilization by specific G4 binders is responsible of a decreased transcription and thus of a reduced gene expression. *C-MYC* G4 behaves as a repressor element while a specific G→T mutation, which destabilize *c-MYC* G4 in the promoter, is associated with an increased *c-MYC* transcription [189].

***KRAS***

*KRAS* promoter contains a high G-content NHE from -327 to -296, positions relative to the exon φ/intron 1 boundary [192]. The first evidence of G4 folding in *KRAS* promoter region was provided by a primer extension assay of a plasmid bearing the template strand. The polymerase arrest was dependent on $K^+$, the cation which better stabilizes G4 conformations. Circular dichroism (CD) and NMR of a 32-mer sequence (32R) confirmed an intramolecular parallel G-quadruplex folding. Two alternative G4 structures, adopted by the 32R (Q1 and Q2), are supported by two independent studies [193, 194]. The 32R has been used as a bait for a pull-down assay with the nuclear pancreatic protein extract. The 32R G4 is recognized and bound by three different proteins: PARP-1, Ku-70 and hnRNP A1 identified through mass spectrometry analysis [193]. HnRNP A1 is one of the most abundant eukaryotic proteins and it is involved in many biological processes as gene expression, nuclear export, RNA splicing and biogenesis [195]. This protein recognizes and unwinds the 32R G4 in the *KRAS* promoter, supporting a biological role in *KRAS* transcription regulation [119]. HnRNP A1 seems to have a general DNA resolvase activity because it is able to unwind several G4 structures present in the genome [196, 197]. A recent study proved a link between *KRAS* expression, ILK and hnRNPA1. The existence of a KRAS-ILK-hnRNP A1 regulatory loop in pancreatic cancer supports the regulating function of hnRNP A1 on *KRAS* expression, through G-quadruplex unfolding [198]. Furthermore, it has been studied also a shorter G4 motif of *KRAS* NHE (21R) and combining different techniques G4 folding has been confirmed also in the 21-mer sequence [199]. Recently, NMR analysis of the 22-mer G4 motif performed by Salgado's group achieved new proofs of G4 folding *in vitro* [200]. If something is known about the G4 role in *KRAS* transcription, even less is known about the complementary C-rich strand and its function. Anyway, Hurley and co-workers have been lately described a possible G-quadruplex/*i*-motif switch involved in *KRAS* expression [201].

***c-KIT***

The *c-KIT* oncogene encodes for tyrosine kinase receptor responsible of cell proliferation, differentiation, apoptosis, chemotaxis and cell adhesion [202, 203]. Aberrant mutations and

overexpression of kit protein results in oncogenic transformation in mast cells, myeloid cell, melanocytes and germ cells. It has a central role in gastrointestinal stromal tumor (GIST) initiation, thus it has been identified as the main target for GIST therapies [204, 205]. Imatinib (*Novartis*) [206] and sunitinib (*Pfizer*) [207] are two recent multitarget kinase inhibitors used in the treatment of *KIT*-positive GIST, because they are active on this receptor. Anyway, drug resistance recurrence [208] create the need of developing a new and different way to target *c-KIT* expression. Also in this case, the presence of three G-quadruplexes (kit1, kit2 and kit*) in the promoter can be seen as an encouraging and attractive way to target *c-KIT* aberrant expression. Kit1 occurs between position -87 and -109 base pairs and kit2 between -140 and -160 bases upstream the TSS and their structures are the best studied. The high-resolution structure of both kit1 and kit2 are available [209-212]. Moreover, the portion comprising the kit1 and kit* motif displays validated binding sites for specific proteins: SP1 (kit1) [213] and AP2 (kit*) [214] indicating again a biological function of these structures. A potential cross talk between the three G4s has been recently suggested by R. Rigo and C. Sissi. They provided evidence for a high order rearrangement of kit2 and kit* into a new intermolecular conformation compatible with a G4-G4 interaction in physiological conditions [215]. Many efforts are being made to understand the complex mechanism of G4s present in the *c-KIT* promoter to put the basis for an efficacious anti-*c-KIT* therapy. During last years, the attention of many scientists has been focused in testing classes of molecules that through G4 binding are able to decrease *c-KIT* transcription. Many *in vitro* studies demonstrated the G4s inhibiting function on *c-KIT* transcription. The negative regulatory role is proved by ligands that bind G4 conformations, increasing their thermal stabilities, induce a decrease of *c-KIT* transcription. *C-KIT* inhibited transcription is associated to an antiproliferative effect [216-218].

***RNA G-quadruplex and translation***

Since 1962, many efforts have been done to understand DNA G-quadruplex structure and its biological function; but surely, RNA G-quadruplexes (RG4) have received less attention during these years. The central core of both DNA and RNA structure is the same, but the assumption that DNA G4 are like their RNA counterparts is certainly an oversimplification. The main difference between DNA and RNA is the presence of the deoxyribose instead of a ribose sugar. The 2'-hydroxyl group in the ribose sugar have several consequences in the RG4 structure. First, it allows many intramolecular interactions within RG4, conferring an enhanced stability. Second, the presence of the hydroxyl group in the grooves leads to an increased capacity to bring water molecules resulting in a more stable conformation [219]. Finally, the ribose sugar influences the RG4 topology. The steric hindrance of the hydroxyl group impairs the *syn*-conformation of the bases thus promoting the *anti*-configuration. Consistently, RNA G-quadruplex is limited to a parallel conformation where all four strands are oriented in the same direction, differently from DNA that can adopt parallel and antiparallel conformations. Recently, it has been also demonstrated that RG4 are more stable than their DNA counterparts under the same buffer conditions [220]. Some authors explored the effect of metal ions on formation and stability of two well-known G4 RNAs: TERRA and NRAS. In a systematic study, they demonstrated that there are significant differences between RG4s and G4 DNAs: RG4s fold into stable quadruplex structure even in buffer only. The study showed that RG4s are less sensitive to metal ions than G4 DNAs. Considering the higher stability of RG4 and that RNA is a single strand nucleic acid extremely "sticky" for its attitude to form secondary structures

(hairpin, loops, bulges etc.) it has been proposed the G-quadruplex folding in mRNA [221, 222], miRNAs [223] and long-non coding RNA (lnRNA)[223]. Computational analysis confirmed the putative RG4 (pG4) folding in 5'- and 3'-UTR (untranslated regions), critical regions for translation regulation [222]. During last years, synthetic RNA oligonucleotides, derived from putative RG4s, have been studied by using structural and biophysical techniques (such as CD, UV melting curves and NMR) and their *in vitro* folding has been disclosed. Moreover, enzymatic and/or chemical footprinting have been largely used to reveal RG4s in both single RNAs and transcriptome-wide studies [223-226]. Important progress in this field achieved from the development of high-throughput RNA sequencing and chemical probing methods allowing RG4s mapping throughout entire transcriptome [225, 226]. Balasubramanian and coworkers developed a RG4 sequencing (rG4-seq) that exploit reverse transcriptase stalling (RTS), caused by the presence of secondary structures as RG4s, coupled with a RNA next generation sequencing. They determined that about 3500 stalls (in approximately 2500 genes) can assume a RG4 conformation [225]. It is worth mentioning that for the transcriptome-wide analysis a purified cellular transcriptome has been used, thus in a cellular context the situation could be different. J. Guo and D. Bartel last year published on Science a substantially different view. They developed a method that combine next generation sequencing sensitivity to selected cations (as $K^+$ vs $Li^+$) and chemical probing techniques, as dimethyl sulfate modification. The core of their work is that the chemical reagents used can penetrate living cells membrane allowing *in vivo* analysis. Surprisingly, they found that RG4s are overwhelmingly not folded and depleted in eukaryotic cells. They additionally showed that RG4 sequences ectopically expressed in Escherichia Coli present in the folded form cause translation and growth inhibition. They explained their results proposing that eukaryotic cells are provided with an unidentified robust machinery responsible for G4 unfolding, differently from bacteria where RG4s are not present in the transcriptome [226]. Although many advances, it is necessary to mention technical limitations in RG4s transcriptome wide analysis. It could be that these methods are not sensible to low levels of folded G4s and so only highly expressed G4s have been measured. Another important consideration: it is possible that G4 folding is a dynamic equilibrium and thus a short-time event (e.g. during splicing or transport); therefore some RG4s are not detected but they have a biological relevance. It is certain that secondary structures in RNA transcripts have a pivotal role in regulating translation [227, 228] so it is intuitive that also RG4 structures probably play a role in this process. Putative G4s enrichment in regulating regions, such as 5'- and 3'-UTRs, supports this notion [222]. In 2001, Moine and collaborators published a glimpse on 5'-UTR RG4 function in translation [229]. Successively, Balasubramanian and colleagues discovered a highly conserved and thermodynamically stable RG4 in the *NRAS* 5'UTR that behaves as repressors in translation regulation [230]. They demonstrated that translation modulation is strictly dependent on the RG4 position in the 5'-UTR. The inhibition is enhanced when the G-quadruplex forming sequence is proximal to the 5' cap [231]. After preliminary studies, many others followed this path. 5'-UTR RG4 are found in several genes (e.g. *TRF2*, *ESR1*, *BCL-2*) reviewed in [232]. The translation inhibition, via RG4 unwinding, is not simply explained by steric impairment of the translational machinery activity; it is also strictly related to G4 distance from cap structure and/or AUG initiation codon [233, 234]. Translation could be regulated by the recruitment of different RBPs (RNA binding proteins) through an enhanced binding site exposure and/or promoting secondary structure. Some studies reported that RG4s within 3'-UTR and ORFs (open reading frame) modulate translation [235, 236]. First evidence that 5'-UTR RG4 inhibits translation has been provided by Arora et al. in 2008 [237]. Furthermore, the clear role of RG4 in orchestrating

protein recruitment led to investigate RG4 possible role in RNA alternative splicing. Genome wide analysis of alternatively spliced transcripts showed more than 3 million of pG4s mapping approximately 30.000 mammalian genes [238, 239]. RG4s near splicing sites can affect and modulate RBPs binding. For example, two RG4s are located near to FMRP-binding site (FBS) on the FMR1 pre-mRNA. FMR1 transcript encodes for two different protein isoforms and FBS is a potent splicing enhancer, which plays a central role in modulating the two spliced isoforms in response to intracellular FMRP levels. Selective mutations, that impaired the RG4 formation, cause a decreased FMRP binding, impair splicing activity of exons and modify splicing pattern of FMR1 pre-mRNA [240]. While in FMR1 case RG4 act as splicing enhancer, the hTERT RG4 behaves as an intronic splicing silencer [241]. These two *in vitro* studies are useful to understand RG4s function in splicing regulation, but surely, the knowledge on this issue has to be improved.

**Proteins binding to G-quadruplex**

G-quadruplex is a non-canonical structure that has been deeply studied in the last years. The biological function of G4 is still not perfectly clear, but its biological relevance is unequivocal. Even if G-quadruplex can act as repressor or as activator/enhancer of a specific biological process, in both cases, its regulatory function is strictly related to proteins that control G-quadruplex folding/unfolding. In literature, it is possible to find a plethora of papers describing how G4s are recognized, bound and unfolded/folded by specific proteins. The proteins involvement is a key point in understanding the complicate mechanism involving G-quadruplex regulatory function in a biological context. *In vitro* studies clearly described how G-quadruplex is a physical hindrance for transcriptional or retro transcriptional enzymes. Indeed, the polymerase stop assay is used as proof of the G-quadruplex folding. *In vitro* polymerases progression is blocked or slowed exactly in G4 proximity. The G4 induce a stop or a transcription pause resulting in a truncated form of the full-length strand [241]. These evidences easily suggest that, in a cellular context, proteins with unwinding activity are required and necessary for transcriptional and/or translational machineries action. Balasubramanian's group last year published a paper [146] describing G-quadruplex correlation with chromatin organization. Strikingly, they pointed out that G4 structures are enriched in promoters and 5'-UTR of highly transcribed genes. They demonstrated that G4 motifs are particularly frequent in cancer-related genes transcriptionally active, as *c-MYC* and *KRAS*, indicating a positive and dynamic relationship between the G4 folding and transcriptional activity. In sight of this novel view, G-quadruplexes can be seen as a sort of beat for protein recruitment. Undoubtedly, the fact that many protein DNA/RNA binding sites are not far from a G4 motif supports this hypothesis [239, 242]. The important role played by proteins is provided also in Bartel's work [225]. They affirmed that eukaryotic cells have been provided by evolution of a protein system involved in the RNA G-quadruplexes unfolding, not present in bacteria. The large amount of proteins that are related with the G4 are collected in the G-quadruplex structure (G4) Interacting Proteins Database (G4IPDB) [243] and in **Table 1** are reported the main proteins involved in *KRAS*, *HRAS* and *c-MYC* oncogenes expression.

| protein name | G-quadruplex | reference |
|---|---|---|
| human-nuclear poly(ADP-ribose)polymerase-1 (h PARP-1) | KRAS (32R) | [119] |
| Ku70 | KRAS (32R) | [119] |
| heterogeneous ribonucleoprotein A1 (hnRNP A1) | KRAS (32R) | [119] |
| Pif1 | c-MYC | [242] |
| Nucleophosmin | c-MYC | [243] |
| Sp1 transcription factor | HRAS | [117] |
| *MYC*-associated zinc finger protein (MAZ) | HRAS | [117] |
| shelterin protein (TRF1,TRF2,POT1,TIN2) | c-MYC | [244] |
| Nucleolin | c-MYC | [245] |
| nucleoside diphosphate kinase 2 (NME/NM23) | c-MYC | [246] |
| DNA methyltransferase 1 (DNMT-1) | c-MYC | [247] |

**Table 1.** Proteins associated to *c-MYC*, *KRAS* and *HRAS* G4s.

**G-quadruplex as therapeutic target**

Since their discovery, G-quadruplex structures have been considered attractive and promising targets in human pathologies, especially in diseases correlated with gene mutations and/or expansions as cancer and neurological disorders. Since aberrant proteins result undruggable like RAS, for years considered a "nightmare for drug developers" [108] targeting a DNA/RNA unusual conformation could be a successful and a promising alternative. Surely, targeting a gene rather than an overexpressed protein or focusing on a specific conformation displayed by an oncogene, means to have few copies of target and thus probably a lower dose of drug. Many oncogenes are fundamental in the tumorigenesis and in the tumor maintenance explaining how cancer cells became "addicted" to them (i.e. RAS, *BRAF*, *c-MYC*). Targeting those oncogenes means to be more selective to tumor cells, as oncogene inhibition will induce apoptosis especially in oncogene-addicted cells. All these advantages explained the grafting work of researchers in designing, screening and validating G-quadruplex ligands. Since the Hurley 'study in 2002 [189], which shows that G4 binders can be used to modulate *c-MYC* expression, many improvements have been achieved in considering G-quadruplex structures as potential new drug targets. A large number of researches highlighted that the majority of G4 binding molecules share common structural features. They are characterized by i) a planar heteroaromatic chromophore necessary for the π–π stacking onto planar G-quartets ii) an extended side chain that is involved in the interaction with G4 grooves and loops and iii) a cationic charge that can promote π-Cation interaction thus enhancing the stacking. Quindoline and berberine derivatives (**Figure 8**) tested in *c-MYC* overexpressing cancer cell lines can be considered classical G4 ligands for their aromatic system linked with one or two amino alkyl side chains [248, 249]. Moreover, the anthraquinone derivatives (**Figure 8**) used in *HRAS*-addicted tumor cells showed high affinity for the G4 in *HRAS* promoter and an anti-proliferative effect in T24 bladder cancer cells [81]. Trisubstituted isoalloxazines are able to bind and stabilize the kit1 and kit2 G4s reducing *c-KIT* mRNA in cells

[216]. The natural macrocyclic telomestatin is an exception as G4 ligand for its chemical structure, but its eight pentagonal rings are sufficient for the π–π stacking interaction with the G-quartets. It has been reported its activity in reducing *c-MYC* expression [250]. A large number of G4 compounds and small molecules have been tested and they exhibited a good G4 binding affinity, but in a drug discovery perspective, an important issue has to be investigated: the G4 ligands selectivity. TMPy4 (**Figure 8**) is a G4 binder that is largely used in the G4-field for its pronounced affinity ($2x10^{-6}$ M) for G-quadruplex structures [251, 252] rather than for duplex conformation. Firstly, selectivity between G4s and other DNA conformation (as DNA duplex) is required and indispensable in a therapeutic view. The duplex is obviously the prevalent conformation in the genome, thus a molecule, that does not discriminate between duplex and G4 structure would have general toxicity and marked side effects, caused by the binding to a large number of genes. Another crucial point for a G4 ligand's selectivity is the ability to recognize a single G-quadruplex among the other present in the whole genome. There are many difficulties to address this issue. It is clear that all G4s share common features such as the central core formed by G-quartets. The grooves are the only structural features that can be exploited to discriminate among G4s. They are formed by the phosphodiester backbones and the loops length determines their dimension, thus G4 grooves are strictly dependent on the sequence and therefore can be unique. Anyway, ligand discrimination remains minimal because it has to overcome global G4 structures. The aimed goal of many studies is to design a unique-G4-specific ligand. A winning solution could be focusing the attention on a specific G4, to identify structural features of that unique G4 to exploit them for a rationale molecular design. The obstacle in this case is the insufficient G4 structural information because less than 1% of the total G4s available in human genome are structurally characterized. The x-ray crystal solution of *c-KIT* G4 in the promoter, for example, permitted to identify a large cleft that seem to be typical only for this structure and is not present in any other known G4. *C-KIT* G4 is a clear and convincing example of how the high-resolution structures and overall structural details can improve ligand design. Trying to find G4 ligands selective solely for a specific conformation could appear a non-realistic challenge, but it is worth mentioning the kinases case. The kinases selectivity years ago seemed to be impossible because of their marked homology and similarity in the ATP binding site, but in 2004, the gefitinib (*Iressa; AstraZeneca*) changed this idea. It is a tyrosine kinase inhibitor selective for the mutated epidermal growth factor receptor (EGFR) expressed in certain human carcinomas. Nowadays this molecule is used in breast and lung cancer therapy [253]. The only G4 ligand who reached the clinical trial (ClinicalTrials.gov identifier: NCT 00780663) is quarfloxin (also named CX-35439), but unfortunately it has not passed phase II. It is a fluoroquinolone derivative chemically optimized to enhance its binding capacity and G4 interaction. Quarfloxin is accumulated in the nucleolus and selectively inhibit Pol I transcription [254]. It disrupts DNA G-quadruplexes-nucleolin complexes, competing with nucleolin G4 binding. The consequent nucleolin release in the nucleoplasm is a common response to cellular stress and activates apoptosis through different pathways [187]. Despite quarfloxin failure in phase II for bioavailability issue, the toxicity profile of this molecule was encouraging. It was well tolerated by patients and it does not display a genotoxic profile. Recently, another compound (CX-5461) entered in early phase of clinical trial for the treatment of breast cancer deficient in BRCA1 and/or BRCA2. Like quorfloxin, it does not cause toxicity and it showed a prominent anticancer activity in experimental animal model [255]. The discovery of new G4 ligands and their validation as therapeutic molecules is the aim of many scientists around the world and recent progress in this field, strongly encouraged efforts in this direction.

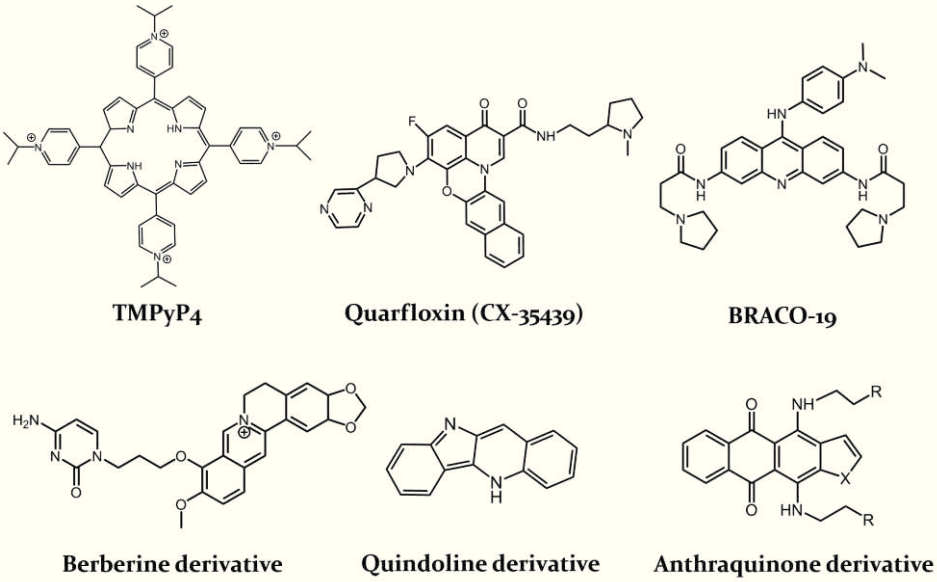



**Figure 8.** G-quadruplex ligands.

### 2.2.2 *i*-motif structure of C-rich strands

In genomic DNA, wherever are present G-rich strands able to form a G-quadruplex structure, there are complementary strands rich in cytosine (C-rich strand). In 1993, Leroy and his coworkers demonstrated for the first time the ability of a DNA C-rich strand to form quadruple-helical structures under acidic conditions [256]. This unusual structure has been named intercalated motif (*i*M) and it consists in a conformation formed by two parallel duplexes present in an antiparallel mode held together through intercalated hemiprotonated cytosine-cytosine base pairs ($C\text{-}C^+$). The cytosine base pair is possible only if the cytosine N3 is present in the protonated form, explaining the strong pH-dependency of the *i*-motif structure. The *i*M folding/unfolding is strictly influenced by pH and this permitted its application in different fields. In nanotechnology, the *i*M is useful in designing nanomachines [257, 258] or in providing the correct assembly of gold nanoparticles [259, 260]. In a biological context, it could be used as a pH sensor to map pH changes in living cells [261, 262] and furthermore it has been used as switch for logic operations [263, 264].

#### *i*-motif topology

The hemiprotonated cytosine base pair has been identified in 1962 [265] and the hairpin formed by $C\text{-}C^+$ base pairs [266] was already known when Leroy and Guèron proposed the *i*-motif conformation. They described the *i*-motif conformation as two cytosinic strands that formed a parallel duplex through $C\text{-}C^+$ base pairs and two of these duplexes again through $C\text{-}C^+$ pairing resulted in an intercalated conformation (**Figure 9**).

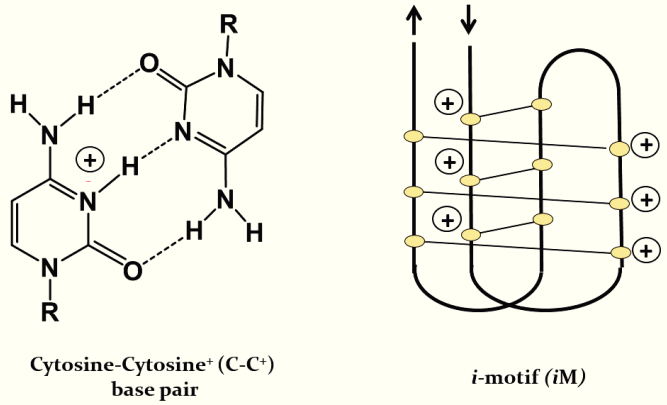



**Figure 9.** Cytosine-Cytosine base pair and *i*-motif structure.

Oligonucleotides can have different intercalation topologies: terminal cytidine can be either on the 3' end or on the 5' end of the strand, respectively classified as 3'E and 5'E topologies [267, 268]. Structures with different intercalation topologies can have similar stabilities. The third possible and less stable topology is the one with a mixture of two or more structures (called the T-form). PH range from 4 to 7 provides the N3 protonation and DNA folding into the closed *i*-motif structure; the highest stability of iM structures occurs at pH value close to cytosine pKa [269]. At higher values, cytosine is deprotonated and it results in the *i*M conversion to the single-stranded form. On the contrary, if the pH is too low (approximately < 3) all the cytosines are protonated not permitting the hydrogen pattern needed for the $C\text{-}C^+$ base pairing [270]. Sequences $d(TAACCC)_n$ present in the human telomeres are also stabilized by interactions between non-cytosine bases. In this case the TAA part can form an A-T base pair which stacks on the cytosine core [269]. Loop interactions are important in determing *i*M stability: in the human telomeric DNA, non-cytosinic bonds are responsible of a thermal stability increase. The systematic substitution of loop bases resulted in a decrease of melting temperature of ~2°C /substitution. The *i*-motif structure is the spatial arrangement of $C\text{-}C^+$ base pairing: if cytosines involved in the structure are all present in one nucleic acid strand it is defined an intramolecular *i*M, on the other hand if it is composed by 2 or many strands it is defined as an intermolecular *i*M. In 2010 Brooks, Kendrick and Hurley categorized the intramolecular *i*Ms into two "classes" depending on the *i*M loop length [271]. "Class I" *i*-motif are characterized by short loop whereas "class II" is formed by *i*M structures with longer loops. "Class II" *i*-motifs are reasonably more stable than "class I" *i*Ms for additional interactions within long loop regions. Moreover, in addition to the three hydrogen bonds at each $C\text{-}C^+$ base pairing, extra interactions in the loop region and $C^+$-positive charges on the backbone favor the four strand association. Therefore, a balance between the sugar interactions and any potential extra interactions in the loops region generates a stable *i*M conformation [272].

### *i*-motif in a biological context

Bioinformatics tools to map *i*-motif structure in the genome has not yet been developed, but starting from the point that the *i*M is naturally the G-quadruplex complementary strand, an overlap in their prevalence in the genome is possible, at least partially. The *i*M structure is particularly

stable at pH values below the physiological one (~7.4) and that is the main reason why it has been considered less interesting and attractive in a biological approach, if compared to the G-quadruplex. Anyhow, recent progress and efforts have been done in understanding how this structure could be possible *in vivo* and in pointing out its biological significance. Special attention has been devoted to the *i*M formation and stability in order to propose an *i*-motif biological role *in vivo*. Obviously, the canonical Watson-Crick duplex in physiological conditions of pH and temperature is the prevalent DNA conformation due to its high thermodynamic stability. Negative superhelicity induced by nuclear processes as transcription, replication and repair, results in a local unfolding of the double helix. In this condition unusual DNA secondary structures as G-quadruplex and *i*-motif could be favored [273]. The importance of supercoiling in modulating the *c-MYC* G4/*i*M folding has been described by Hurley and colleagues [274]. They employed a system to induce negative supercoiling upstream the TSS in a supercoiled plasmid. They inserted into a Del4 plasmid wild-type or mutated sequence of the *c-MYC* NHE III and comparing chemical/enzymatic footprintings they showed the negative superhelicity contribution in promoting secondary DNA structures at physiological conditions [274]. A cancer hallmark is certainly an altered metabolism and one of the main consequences is surely a dysregulated pH. In normal and differentiated cells the intracellular pH is generally around 7.2 and the extracellular one is ~7.4. Differently, a typical feature of a cancer cell is a higher intracellular pH (~7.4) and a lower extracellular pH (6.7-7.1) [275, 276]. The acidic environment of cancer cells could be a biological condition in which the *i*M folding could be promoted and favored resulting in an altered gene expression. Moreover, some biological processes could cause a local and transient acidification in the cell. For example, Poly (ADP-ribose) polymerase (PARPs) produces 1 mol of proton nicotinamide for each mol of NAD consumed. Therefore, PARPs activation is responsible for protons release in a very short period resulting in a transient acidification [277], which could drive and regulate *i*-motif folding in cells. Another important consideration must be the formation of the *i*M conformation in a biological environment: the crowding nature of the *in vivo* condition. The cellular context is surely characterized by a crowded environment defined as the excluded volume effect and dehydration effect. Rajendran and colleagues demonstrated that polyethylene glycol (PEG 2000 and PEG 8000), used to mimic the molecular crowding *in vitro*, induces the *i*M folding at neutral pH value (pH 7.0) [278]. These outfindings have been showed also for the *c-MYC* promoter sequence: *c-MYC i*-motif forms a stable structure at pH values as high as 6.7 in 40% w/w polyethylene glycol with a molecular weight higher than 12000 g/mol [279]. In spite of the uncertain existence *in vivo* it has been proved that C-rich sequences, so putative *i*-motif conformations, are source of error in DNA amplification [280]. The C-rich region of the murine *KRAS* promoter is one of the first oncogene *i*-motif that has been characterized [281]. In the last years, several *i*Ms present in oncogene promoters have been investigated and described such as *c-MYC* [282], *BCL2* [283], *VEGF* [284], *RET* [285], *Rb* [286]. In addition, a dual *i*-motif/G-quadruplex system in *KRAS* promoter has been presented as a possible mechanism to modulate gene expression [201]. The increasing number of studies on this topic in the last years demonstrates a growing interest on the *i*M structure and on its possible involvement in gene expression regulation.

**Proteins recognizing the *i*-motif**

In order to support the biological relevance of the *i*-motif folding, protein interacting with C-rich DNA strands have been deeply investigated. The *i*M geometry and charge distribution could be a

sophisticated mode to enhance specific structural protein recruitment [287]. Hurley's group presented a detailed mechanism in which transcription factor hnRNP LL activates *BCL-2* transcription. Through electrophoretic mobility shift assays (EMSA) the study shows that the *i*-motif folding is needed for hnRNP LL binding. The hnRNP LL binding does not occur if the protein is incubated with a mutated C-rich sequence incapable of forming the *i*M structure [288]. Hurley's *in vitro* study is the first example of how gene expression could be modulated targeting the *i*-motif structure. They demonstrated that *BCL-2* promoter exist in a dynamic equilibrium between *i*M and hairpin conformations. Through the addiction of two compounds, one selective for the *BCL-2-i*-motif (*IMC-48*) and the other for the hairpin conformation (*IMC-76*), the equilibrium is shifted to each structure, respectively. *I*-motif stabilization lead to a *BCL-2* up-regulated expression and on the other hand the hairpin impairs the *i*M structure and causes a downregulated *BCL-2* transcription. Furthermore, hnRNP LL transcription factor, through RNA recognition motifs (RRM) 1 and 2, is able to recognize and bind the *i*-motif loop II and loop VI promoting a conformational change which yield *BCL-2* transcriptional activation [288]. The discovery of the *i*-motif binding protein hnRNP LL supports, in the BCL-2 case, a plausible role of this unusual DNA structure in a biological process. The *i*M hypothesized function is to provide a structural arrangement for protein recognition causing the unfolding and so the transcription initiation. They proposed the *i*-motif as a mode to display specific nucleotide sequences (loops region) in the most kinetically favorable conformation for protein binding. The protein recognition yields the structural unwinding and promotes single stranded form, needed for a "fluent" transcription. HnRNP K is a transcription factor, belonging to the same protein family able to bind the C-rich region present in the *c-MYC* promoter [289]. It has similar spaced domains that recognize the *c-MYC* TCCC sequences [290]. As in *BCL-2* promoter, the recognition sequences are present in the lateral loops of the *c-MYC i*-motif and are spaced by the same number of bases, hence, hnRNP K may have the analogous regulatory function on *c-MYC* transcription. Additionally, it has been shown that hnRNP K is able to bind the C-rich strand of human telomeres that displays the same CCCT repeats [287]. Despite the fact that these sequences of the telomeric *i*-motif are not present in the lateral loops, Lacroix and coworkers showed that protein binding to the human telomeric *i*-motif is stable at pH values ≥ 6 [287]. Maurice Gueròn's laboratory investigated the proteins binding to the yeast, *tetrahymena* and vertebrate telomeric sequences that are expected to form *i*-motif structures [291]. In literature a large variety of protein that selectively bind C-rich single strand are reported [292, 293] and it could be possible that their recruitment depends on a secondary DNA structure. Further studies in this direction are necessary to better understand the *i*-motif physiological role. The perception of G-quadruplex has changed in time, from a simple unusual DNA structure to a potential therapeutic target involved in many oncogenes expression. Compared to the deep knowledge on G-quadruplex, our information on the *i*-motif structure and its biological implications needs to be improved. Anyway, the growing evidences supporting the *i*M folding in physiological conditions can encourage researches and stimulate scientists in proposing the *i*-motif structure as hypothetical new potential target for genetic diseases.

# 2. AIM OF MY PhD WORK

As the ras genes are considered "undruggable" since a long time, the search for alternative anti-cancer strategies to suppress their expression in ras addicted tumors is one of the most important goals to be pursued in cancer therapy. My PhD work started from the above consideration; I have investigated the role of unusual nucleic acid structures in the ras promoters and in the 5'-untraslated region (5'-UTR) of KRAS with the aim of developing alternative anti-ras strategies. Specifically, my work focused on the regulatory function of DNA and RNA non-canonical structures (G-quadruplex and *i*-motif) associated to the *HRAS* promoter and to the KRAS 5'-UTR (**Figure 10**). In addition, I have developed a single-stranded miRNA-oligonucleotide therapy specific for oncogenic *KRAS*.

In the first part of my research, I focused on the *HRAS* oncogene. It is often mutated in bladder tumours and since it is involved in the pathogenesis of this disease, its downregulation is expected to sensitize cancer cell to chemotherapy. Previous studies made in our laboratory demonstrated that the *HRAS* oncogene is regulated at transcriptional level by a molecular switch involving two G-quadruplex structures located immediately upstream of the transcription start site [117]. Considering that the formation of G4 DNA by the G-rich strand leaves the complementary C-rich strand unpaired and subjected to nuclease, we asked if the latter is able to fold into the so called i-motif conformation. I therefore focused on the C-rich complementary strand of the *HRAS* promoter. I have investigated the capacity of the HRAS C-rich strand to fold into the *i*-motif conformation at several conditions, including a near-physiological condition. I also discovered that hnRNP A1, a protein associated to the CG-elements of the ras genes, is able to bind to the i-motif and unfold this particular DNA structure. On the basis of this finding we proposed the formation in the HRAS promoter of a G-quadruplex/*i*-motif switch controlling transcription in bladder cancer cells. Subsequently, as the *HRAS* G-quadruplex is recognized by specific transcription factors (Sp1, MAZ and hnRNP A1) essential for transcription, I employed a decoy strategy to suppress *HRAS* in bladder cancer cells. Its rationale is based on the principle that delivering to the cells oligonucleotides that mimic the *HRAS* G-quadruplexes, they will compete with the binding of the transcription factors to the promoter with the result that transcription will be decreased, if not completely suppressed. We designed decoy oligonucleotides engineered with anthraquinone insertions and locked nucleic acids (LNA) modifications, mimicking one of the two neighbouring G-quadruplexes controlling HRAS expression.

My PhD work then focused on KRAS which is mutated in pancreatic cancer cells. The aim was to suppress this oncogene through strategies based on the use of small molecules. To address this issues I focused on the transcript of KRAS, especially on the 5'-UTR and 3'-UTR untranslated regions. The 5'-UTR is very rich in guanines (77 % CG) and contains three non overlapping G-quadruplex motifs in the first 80 nt. Using a combination of techniques, we demonstrated that the 5'-UTR exists in equilibrium between a stem-loop structure and a G-quadruplex structure. I used a new class of small molecules (anthrathiophenediones and anthrafuranediones) to target the G-quadruplexes of the KRAS 5'-UTR. With my work I found that these small molecules are able to inhibit the process of translation. The suppression of KRAS induced apoptosis in pancreatic cancer cells. I also focused on the 3'-UTR, where there is the target region of miR-216b: a microRNA that is aberrantly down-regulated in pancreatic cancer. MiRNAs are small non-coding RNA endogenous

molecules (containing about 22 nucleotides) that function as post-transcriptional regulators of gene expression. We designed single-stranded miR-216b mimics, modified with unlocked nucleic acid (UNA) insertions to increase their nuclease resistance. We evaluated the efficacy of a new delivery strategy for miR-216b mimics based on the use of palmitoyl-oleyl-phosphatidylcholine (POPC) liposomes functionalized with lipid-modified miR-216b and lipid-modified cell penetrating TAT peptide.
The results obtained by our work have been published and/or submitted in peer-reviewed journals and are reported in the "Results" section of this dissertation.

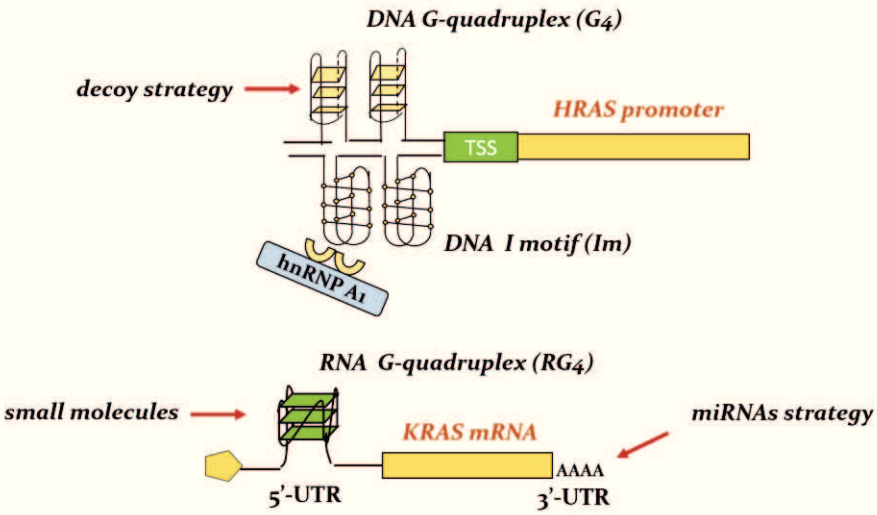



**Figure 10.** Cartoon to recap the aim of the study.

## List of articles:

- **Section A**
  G. Miglietta, S. Cogoi, EB. Pedersen, LE. Xodo, **GC-elements controlling HRAS transcription form *i*-motif structures unfolded by heterogeneous ribonucleoprotein particle A1,** Scientific reports 5 (2015)
- **Section B**
  G. Miglietta, AS. Gouda, S. Cogoi, EB. Pedersen, LE. Xodo, **Nucleic acid targeted therapy: G4 oligonucleotides downregulate HRAS in bladder cancer cells through a decoy mechanism**, ACS medicinal chemistry letters 6 (12), 1179-1183 (2015)
- **Section C**
  G. Miglietta, S. Cogoi, J. Marinello, G. Capranico, AS. Tikhomirov, A. Shchekotikhin, LE. Xodo, **RNA G-quadruplexes in Kirsten ras (KRAS) oncogene as targets for small molecules inhibiting translation**, Journal of Medicinal Chemistry *60* (23), 9448–9461 (2017)
- **Section D**
  G. Miglietta, A. Ferino, S. Cogoi, R. Picco, S. Vogel, J. Wengel, LE. Xodo, **Potent anti-Kirsten ras (KRAS) single-stranded miRNA oligonucleotide mimics as therapeutic agents**, in preparation.

**Table representing my personal contribution in the published papers.**

| paper | Project planning | Experimental part | Data analysis | Manuscript preparation |
|---|---|---|---|---|
| GC-elements controlling HRAS transcription form i-motif structures unfolded by heterogeneous ribonucleoprotein particle A1. (**Section A**) | **5%** | **70%** | **20%** | **10%** |
| G4 Oligonucleotides downregulate HRAS in bladder cancer cells through a decoy mechanism. (**Section B**) | **10%** | **85%** | **30%** | **10%** |
| RNA G-quadruplexes in Kirsten ras (KRAS) oncogene as targets for small molecules inhibiting translation. (**Section C**) | **50%** | **100%** | **50%** | **40%** |
| Potent anti Kirsten ras (KRAS) single-stranded miRNA oligonucleotide mimics as therapeutic agents. (**Section D**) | **20%** | **80%** | **50%** | **20%** |

## 4. RESULTS

**Section A:**

# GC-elements controlling HRAS transcription form *i*-motif structures unfolded by heterogeneous ribonucleoprotein particle A1

G. Miglietta, S. Cogoi, EB. Pedersen and LE. Xodo

Scientific reports 5, 2015

SCIENTIFIC REPORTS

OPEN

# GC-elements controlling *HRAS* transcription form *i*-motif structures unfolded by heterogeneous ribonucleoprotein particle A1

Received: 19 August 2015
Accepted: 11 November 2015
Published: 17 December 2015

Giulia Miglietta[1,*], Susanna Cogoi[1,*], Erik B. Pedersen[2] & Luigi E. Xodo[1]

***HRAS*** **is regulated by two neighbouring quadruplex-forming GC-elements (*hras*-1 and *hras*-2), located upstream of the major transcription start sites (doi: 10.1093/nar/gku 5784). In this study we demonstrate that the C-rich strands of *hras*-1 and *hras*-2 fold into *i*-motif conformations (*i*Ms) characterized under crowding conditions (PEG-300, 40% w/v) by semi-transitions at pH 6.3 and 6.7, respectively. Nondenaturing PAGE shows that the *HRAS* C-rich sequences migrate at both pH 5 and 7 as folded intramolecular structures. Chromatin immunoprecipitation shows that hnRNP A1 is associated under *in vivo* conditions to the GC-elements, while EMSA proves that hnRNP A1 binds tightly to the *i*Ms. FRET and CD show that hnRNP A1 unfolds the *i*M structures upon binding. Furthermore, when hnRNP A1 is knocked out in T24 bladder cancer cells by a specific shRNA, the *HRAS* transcript level drops to 44 ± 5% of the control, suggesting that hnRNP A1 is necessary for gene activation. The sequestration by decoy oligonucleotides of the proteins (hnRNP A1 and others) binding to the *HRAS* *i*Ms causes a significant inhibition of *HRAS* transcription. All these outcomes suggest that *HRAS* is regulated by a G-quadruplex/*i*-motif switch interacting with proteins that recognize non B-DNA conformations.**

The *HRAS* oncogene encodes for a 21-k*D* GTP-ase conveying signals to the nucleus that stimulate cell proliferation[1]. In many tumours *HRAS* is mutated, normally in exon 1, codon 12, 13 or 61, and encodes for an altered protein which constitutively activates downstream pathways causing normal cells to become cancerous cells[2]. In previous works, we have demonstrated that *HRAS* is regulated by two neighbouring GC-rich elements that we called *hras*-1 (nt 432-464, A.N. J00277) and *hras*-2 (nt 509-530, A.N. J00277), located immediately upstream of the major transcription start sites (TSS's), each capable of folding into a G-quadruplex structure[3,4]. By site-directed mutagenesis of the GC-elements, we found that the G-quadruplexes behave as transcription repressors[3]. Under normal conditions, *hras*-1 and *hras*-2 are folded into G-quadruplexes, thus locking the promoter into an inactive state characterized by a low transcription level[3]. Transcription is activated when the G-quadruplexes are unfolded, and the G-elements transformed into canonical B-DNA forms. We found that MAZ, a zinc-finger transcription factor recognizing blocks of guanines, interacts with the promoter GC-elements under cellular conditions[3]. MAZ is an essential protein for gene expression, as it unfolds the *HRAS* G-quadruplexes and activates transcription[3,4]. Our data support a transcription model according to which the two neighbouring G-quadruplexes behave as a molecular switch that controls gene expression.

In the present work we interrogated if the complementary C-rich strands of *hras*-1 and *hras*-2 (namely *hras*-$1^{Y}$ and *hras*-$2^{Y}$) fold into the well known *i*M conformation[5–15]. We found that *hras*-$1^{Y}$ and *hras*-$2^{Y}$ assume the *i*M conformation under slightly acidic conditions, which are close to neutrality in the presence of a crowding agent, for example PEG-300[16]. We also discovered that the *HRAS* *i*Ms are recognized by nuclear proteins, including nuclear factor hnRNP A1. This protein, which shows a binding preference for cytosines, unfolds the *i*M conformation of

[1]Department of Medical and Biological Sciences, P.le Kolbe 4, 33100 Udine, Italy. [2]Nucleic Acid Center, Institute of Physics and Chemistry, University of Southern Denmark, DK-5230 Odense M, Denmark. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to L.E.X. (email: luigi.xodo@uniud.it)

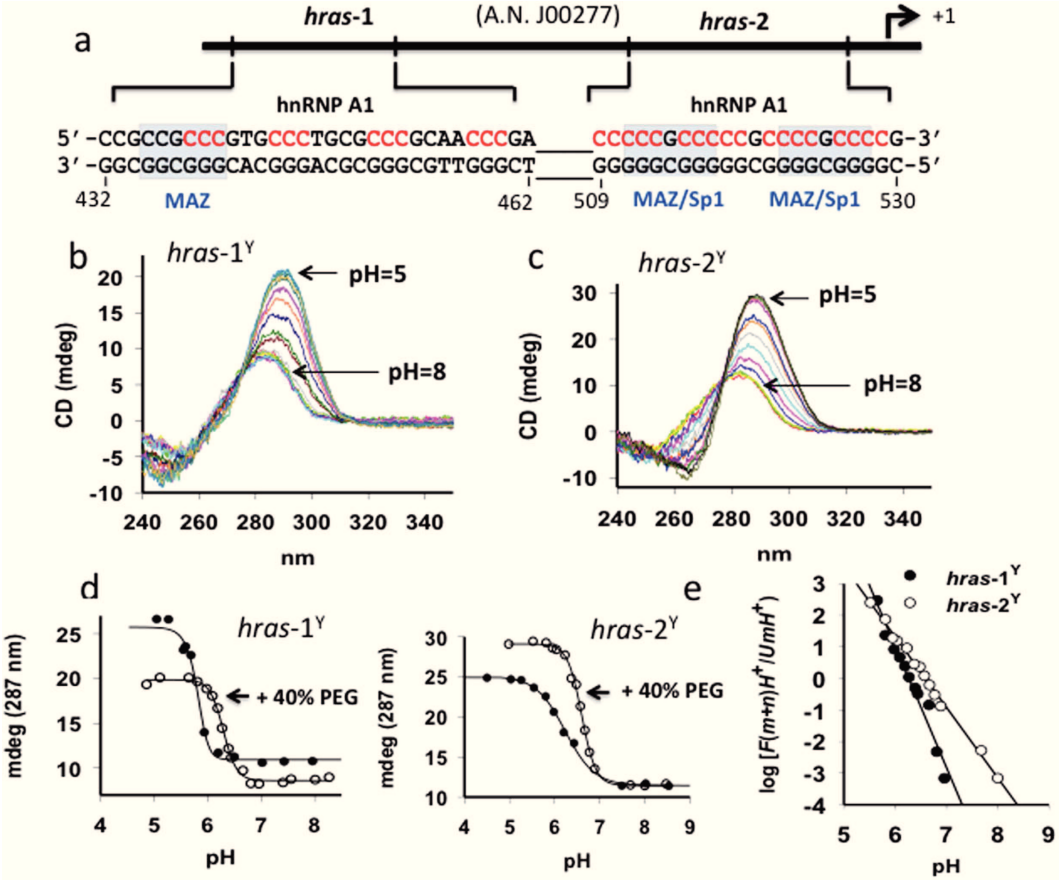



**Figure 1.** (**A**) Sequences of the GC-rich elements located in the *HRAS* promoter upstream of major TSS's; (**B,C**) Circular dichroism titrations of *hras*-1$^{Y}$ and *hras*-2$^{Y}$ (3 μM, 1 cm pathlength cell) in 50 mM Tris-acetate, 50 mM KCl, 40% PEG-300 and pH from 4.5 to 8; (**D**) Ellipticity (287 nm) *versus* pH curves for *hras*-1$^{Y}$ and *hras*-2$^{Y}$ in the presence and absence of PEG-300; (**E**) Determination of number of protons picked up by *hras*-1$^{Y}$ and *hras*-2$^{Y}$ upon folding into the *i*M.

the *HRAS* sequences. When hnRNP A1 was knocked out in T24 bladder cancer cells by a specific shRNA, the level of *HRAS* transcript also dropped to 44 ± 5% of the control. Together, our data provide evidence that hnRNP A1, with its unfolding activity against the *i*M, is an essential factor for the activation of *HRAS*. Indeed, when hnRNP A1 was sequestered by decoy oligonucleotides mimicking the *i*M, *HRAS* transcription was significantly down-regulated. The outcome of this work support the notion that *HRAS* expression is regulated by a G-quadruplex/*i*M switch that is controlled by proteins.

## Results and Discussion

The sequence of the *HRAS* promoter immediately upstream of the major transcription start sites is reported in Fig. 1A. It contains two GC-rich elements, *hras*-1 and *hras*-2, composed of blocks of guanines and capable of folding into G-quadruplex structures. In previous works we have demonstrated that these sequences behave as a regulatory switch controlling gene expression[3,4]. Such a mechanism has been proposed for other relevant oncogenes including *KRAS*[17,18], *CKIT*[19,20], and *CMYC*[21,22]. A couple of comprehensive reviews on this subject have been reported[23,24]. In this work we have focused on the complementary C-rich strands *hras*-1$^{Y}$ and *hras*-2$^{Y}$ and have investigated if they fold into stable *i*Ms.

***i*M formation by the *HRAS* C-rich sequences.** To find out if *hras*-1$^{Y}$ and *hras*-2$^{Y}$ can assume the *i*M conformation, we performed circular dichroism (CD) experiments as a function of pH, in 50 mM KCl, 50 mM Tris-acetate, 25 °C. To mimic the crowding conditions of the cell, we analysed the sequences both in the presence and absence of 40% (w/v) PEG-300[16]. Typical CD titrations are shown in Fig. 1B,C. It can be seen that the spectra of *hras*-1$^{Y}$ and *hras*-2$^{Y}$ change dramatically as the pH is gradually decreased from 8 to 4.5. Under acidic conditions (pH 5) both sequences exhibit the characteristic enhanced ellipticity at ~287 nm of a classical *i*M[10,12,25–27], while at pH 8 the sequences exhibit a much lower ellipticity, shifted at ~285 nm. By plotting the 287-nm ellipticity as a function of pH, we obtained for each sequence, in the presence or absence of PEG-300, sigmoidal curves reflecting *i*M formation (Fig. 1D). The crowding agent drives the folding at higher pH values: the semi-transition of *hras*-1$^{Y}$ increases from pH 5.9 to 6.3, whereas that of *hras*-2$^{Y}$ increases from pH 6.2 to 6.7. These plots suggest that the *i*Ms are stable in a slightly acidic medium. However, under cellular conditions the *i*M can be stabilized by:

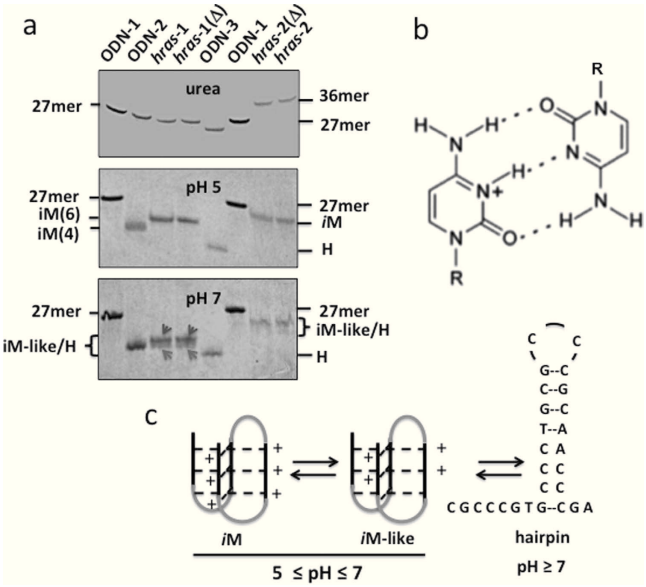



**Figure 2.** (**A**) PAGE of *hras*-1$^{Y}$ and *hras*-2$^{Y}$ and variants ODN-1, ODN-2 and ODN-3 in denaturing conditions (7 M urea); native conditions at pH 5 and 7. The oligonucleotide sequences are reported in Methods. The electrophoresis was run at 20 °C, the bands stained with stains-all. Δ = heated 10 min 90 °C; (**B**) $CH^+$:C base pair; (**C**) Putative equilibrium between *i*Ms involving different protonation levels and a flexible hairpin formed by *hras*-1$^{Y}$ at pH ≥ 7.

(i) transcriptionally induced DNA superhelicity[28,29]; (ii) more effective cellular crowding conditions[30,31]; (iii) an increased intracellular acidity generated by an increase of the glucose-lactate flux[32,33].

As the ellipticity-*versus*-pH curves are reversible, we evaluated the number of protons involved in the folding of *hras*-1$^{Y}$ and *hras*-2$^{Y}$, by considering the following equilibria:

$$mH^+ + U \rightleftarrows U^* \cdot mH^+ \quad (1)$$

$$U^* \cdot mH^+ + nH^+ \rightleftarrows F \cdot (m+n)H^+ \quad (2)$$

where equilibrium 1 takes into consideration the fact that the unfolded C-rich sequence, *U*, may bind *m* protons before folding into the *i*-motif (as the pKa of cytosine is ~4.4[34], at pH 5 about 3 cytosines out of 12 are expected to be protonated); $U^*{\cdot}mH^+$ is the unfolded sequence with *m* bound protons. The bases in $U^*{\cdot}mH^+$ are assumed to be unstacked, so the formation of this species is accompanied by a negligible change in the CD spectrum[35]. In contrast, upon folding into *i*-motif *F*, $U^*{\cdot}mH^+$ assumes in a cooperative manner other *n* protons to form (*n*+*m*) $CH^+$:C base pairs that are stacked in the structure. A significant CD spectral change is expected for the formation of the *i*M[8,10] (equilibrium 2), whose equilibrium constant is $K_D = ([F{\cdot}(m+n)H^+]/([U^*{\cdot}mH^+][H^+]^n)$ (3). From the CD plots of Fig. 1B,C, we determined for each sequence the fraction of folded and unfolded species and ratio $[F(n+m)H^+]/[U^*{\cdot}mH^+]$. By plotting log $[F(n+m)H^+]/[U^*{\cdot}mH^+]$ *versus* pH, we obtained a straight line whose slope is *n* (Fig. 1E). We obtained values of $n \sim 4$ and $n \sim 2$ for *hras*-1$^{Y}$ and *hras*-2$^{Y}$, respectively. This suggests that when *hras*-1$^{Y}$ and *hras*-2$^{Y}$ fold into the *i*M, they assume 4 and 2 protons, respectively, which agrees with the fact that the sequences are partially protonated before folding. A similar behaviour has been previously observed for the formation of the *i*-motif by $(C_3TA_2)_4$[14].

**PAGE and analyses of melting curves.** The intramolecular *i*M, being a folded structure, migrates in a polyacrylamide gel faster than an unfolded oligonucleotide of the same length. We analysed *hras*-1$^{Y}$ and *hras*-2$^{Y}$ by PAGE under different pH conditions. The mobility of the two C-rich sequences was compared with that of *hras*-1$^{Y}$ variants: ODN-1 (unable to form any structure); ODN-2 (forming an *i*M with 4 $CH^+$:C) and ODN-3 (forming a stable W.C. hairpin) (Fig. 2A,B) (Methods). Under denaturing conditions (7 M urea), the 27-mer oligonucleotides (*hras*-1$^{Y}$ and variants) exhibited the same mobility, with the exception of ODN-3 that forms a hairpin even in the presence of 7 M urea. Sequence *hras*-2$^{Y}$, being embedded in a 36-mer oligonucleotide, migrates slowly (Methods). In contrast, under native conditions at pH 5, *hras*-1$^{Y}$, that assumes the *i*M according to CD, migrates with a sharp band faster than that of ODN-1. Interestingly, when the *hras*-1$^{Y}$ sequence is modified to fold into a hairpin stabilized by a stem of 8 W.C. bps (ODN-3) ($S_1$), it migrates even quicker than the *i*M. This is because the 6 positive charges of *i*M reduce its negative charge density. The folding of *hras*-1$^{Y}$ into the *i*M is quite fast, as the mobility does not change when the sample is heated before loading. Variant ODN-2, forming an *i*M with 4 positive charges ($S_1$), migrates slightly faster than *hras*-1$^{Y}$, as expected. At pH 5, the 36-mer oligonucleotide containing sequence *hras*-2$^{Y}$ migrates quicker than unstructured 27-mer ODN-1, as it folds into the *i*M (see CD). At pH 7, both *hras*-1$^{Y}$ and *hras*-2$^{Y}$ still migrate faster than ODN-1, indicating that even at neutral pH the sequences are folded. However, they

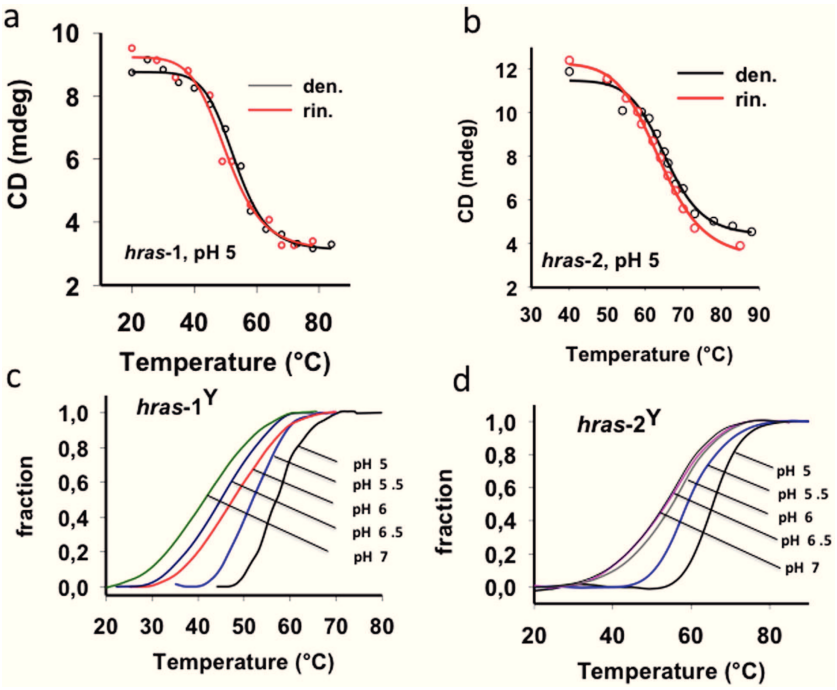



**Figure 3.** (**A,B**) CD melting curves of 3 μM *hras*-1$^{Y}$ and *hras*-2$^{Y}$ in 50 mM sodium cacodylate, pH 5, 50 mM KCl. Denaturing curve (20 → 85 °C), renaturing curve (85 → 20 °C); (**C,D**) Fraction of *i*M *versus* T curves obtained from FRET-melting experiment (0.3 °C/min), in 50 mM sodium cacodylate, pH 5 to 7, 50 mM KCl. The curves of *hras*-1$^{Y}$ and *hras*-2$^{Y}$ at pH 5 are fully reversible. FRET-melting gives $T_M$ values ~3 °C higher than CD-melting values, due to the presence of the fluorophores in the oligonucleotides analysed by FRET.

migrate either with a smeared band (*hras*-1$^{Y}$) or with two bands (*hras*-2$^{Y}$), suggesting that more than one folded structure is formed: most likely *i*M-like structures stabilized by C:C and CH$^{+}$:C bps. It is also possible that at pH 7, the *i*M-like structure is in equilibrium with a flexible hairpin (Fig. 2C, $S_1$).

We have also examined the thermal stability of the *i*Ms by CD- and FRET-melting experiments. Typical CD-melting profiles for *hras*-1$^{Y}$ and *hras*-2$^{Y}$ are reported in Fig. 3A,B. By heating (20 → 85 °C) and then cooling (85 → 20 °C) the DNA solutions in 50 mM sodium cacodylate, pH 5, 50 mM KCl, the folded/unfolded transitions showed to be cooperative and reversible, as previously found for C-rich oligonucleotides under similar experimental conditions[10,12–14]. The melting in the pH range between 5 and 7 was examined by FRET experiments, using oligonucleotides end-labelled with ATTO-488 (5′-end) and TAMRA (3′-end) (Fig. 3C,D). The $T_M$ of both *hras*-1$^{Y}$ and *hras*-2$^{Y}$ decreased with increasing pH: *hras*-1$^{Y}$, 55.9, 52.1, 47.6, 44.8, 41.3 °C at pH 5, 5.5, 6, 6.5, 7, respectively; *hras*-2$^{Y}$, 65.5, 58.7, 55.1, 53.8, 53.1 °C at pH 5, 5.5, 6, 6.5, 7. At pH values > 5, the $T_M$ of the *i*Ms decreases as the structure is probably stabilized by a number of CH$^{+}$:C < 6. It is indeed reasonable to assume that when the medium is not sufficiently acidic, the *i*M is stabilized by both C:C and CH$^{+}$:C base pairs[36,37]. *i*M structures of different protonation levels may coexist in solution. Mechanical stability experiments with the ILPR C-rich sequence showed that partially folded *i*M-like species are in equilibrium with fully folded *i*M at neutral pH[36]. The higher $T_M$ of *hras*-2$^{Y}$ is likely due to the additional CH$^{+}$:C bp that stabilizes the *i*M. The two structures have similar rupture forces and sufficient stability to stall RNA polymerase[36]. Yang and Rodgers have reported that the energy of C:C is about 1/3 of that of CH$^{+}$:C[38]. Sequence *hras*-2$^{Y}$ shows a behaviour similar to *hras*-1$^{Y}$, with the difference that at pH 5 it shows a higher $T_M$, 65.5 °C (Fig. 3D).

We evaluated the thermodynamics of the folding transitions according to a two-state model. This was done at pH 5, where the two sequences fold into only one structure, as shown by PAGE. From the CD- and FRET-melting curves at pH 5, we obtained the following average thermodynamic data (±10%): $\Delta H = 252$ kJ/mol and $\Delta S = -770$ J/mol K and $\Delta G = -17$ kJ/mol for *hras*-1$^{Y}$; $\Delta H = 323$ kJ/mol and $\Delta S = -950$ J/mol K and $\Delta G = -27$ kJ/mol for *hras*-2$^{Y}$. Assuming that the breaking of a CH$^{+}$:C bp needs approximately $46 \pm 4$ kJ/mol[12], the number of CH$^{+}$:C broken by the thermal disruption of *hras*-1$^{Y}$ is ~6, of *hras*-2$^{Y}$ is ~7, in accord with the number of expected protons that should bind to the sequences at pH near pKa[10,12]. At pH 5, nearly half of the cytosines is protonated and the sequences are completely folded into *i*M, showing the highest stability.

**The *HRAS i*Ms are recognized by hnRNP A1.** As the *i*M-forming sequences overlap critical GC-elements immediately upstream of TSS, we interrogated if these unusual structures are recognized by nuclear proteins. Previous studies have reported that DNA sequences composed by runs of cytosines such as the C-repeats in the telomeres and in the *CMYC* promoter are recognized by proteins of the heterogeneous nuclear riboproteins family (hnRNP)[39,40] Moreover, Hurley and co-workers recently reported that the *i*M formed in the *BCL2* promoter interacts with hnRNP LL[6]. Specific binding of heterogeneous ribonuclear proteins to C-rich DNA sequences is

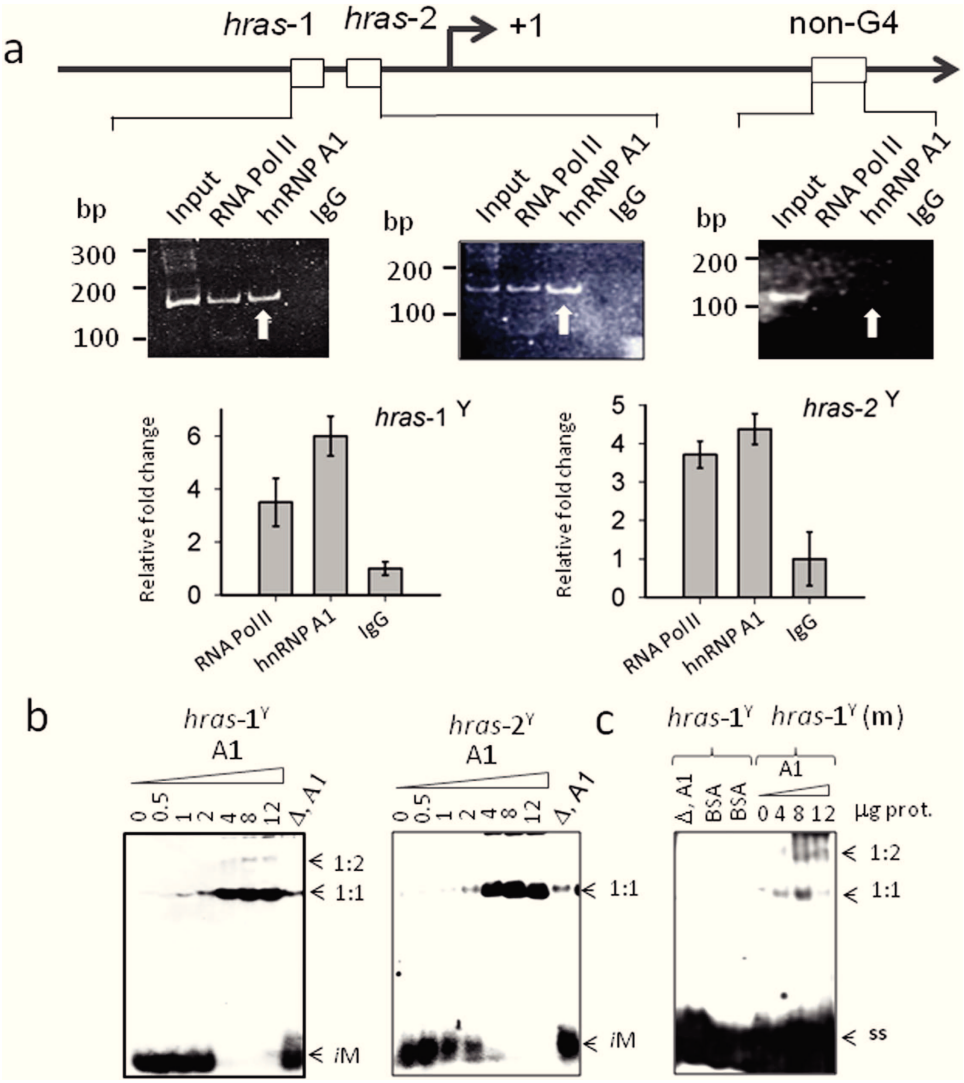



**Figure 4.** (**A**) ChIP experiment to determine the occupancy of *hras*-1, *hras*-2 and control sequence (870 bp downstream from TSS) by hnRNP A1. Histograms shows the relative occupancy of *hras*-1 and *hras*-2 by hnRNP A1, RNA Pol II (positive control) and IgG (negative control). Data have been normalized by IgG signal; (**B**) EMSA of $^{32}$P-labelled *hras*-1$^{Y}$ and *hras*-2$^{Y}$ in 50 mM Tris-acetate pH 5.5, 50 mM KCl, incubated 40 min at room temperature with increasing amounts of recombinant hnRNP A1 (0–12 μg). Lane (Δ,A1) indicates the *i*M incubated 40 min at room temperature, with denatured hnRNPA1 in binding buffer (see Methods); (**C**) EMSA at pH 5.5 of *hras*-1$^{Y}$ with BSA or denatured hnRNP A1 and EMSA of *hras*-1$^{Y}$ (m) with hnRNP A1; ss = single-stranded oligonucleotide; 1:1 and 1:2 DNA-protein complexes.

also supported by previous work, according to which hnRNP A1 binds to the GC-element of *KRAS*[17], which shows a high sequence/functional homology with the *HRAS* GC-elements. HnRNP A1 is one of the most abundant nuclear proteins of eukaryotic cells that regulates several aspects of mRNA biogenesis[41]. As it is over-expressed in a variety of cancers[41,42], we wondered if this protein plays a role in the promoter of the *HRAS* oncogene, in the region where the *i*M can potentially be formed. To address this question, we first investigated by chromatin immunoprecipitation (ChIP) if in *HRAS*-mutant T24 bladder cancer cells, hnRNP A1 is associated to the GC-elements under *in vivo* conditions. The occupancy of *hras*-1 and *hras*-2 (located 6 bp upstream of first TSS) by hnRNP A1 was compared with the occupancy of a reference GC-rich sequence unable to fold into a non-B DNA structure (located 870 bp downstream from first TSS). A typical ChIP is shown in Fig. 4A. We found that the occupancy of *hras*-1 and *hras*-2 by hnRNP A1 was, respectively, ~6- and ~5-fold higher than the occupancy by IgG (negative control). As *hras*-1 and *hras*-2 are located in the region of the major transcription start sites, they show a significant occupancy by RNA Pol II: ~4-fold higher than the IgG signal. In contrast, the reference sequence showed almost no occupancy by any of the proteins considered. The ChIP data provided strong evidence that hnRNP A1, under *in vivo* conditions, is indeed associated to the critical GC-elements of the *HRAS* promoter. However, ChIP data do not provide information about the conformation of the GC-elements interacting with hnRNP A1. To know if the nuclear factor recognizes the *i*M, we performed EMSA at pH 5.5 and 20 °C of mixtures composed by *hras*-1$^{Y}$ or *hras*-2$^{Y}$ and recombinant hnRNP A1, which was produced with a high degree of purity (Fig. 4B,C). It

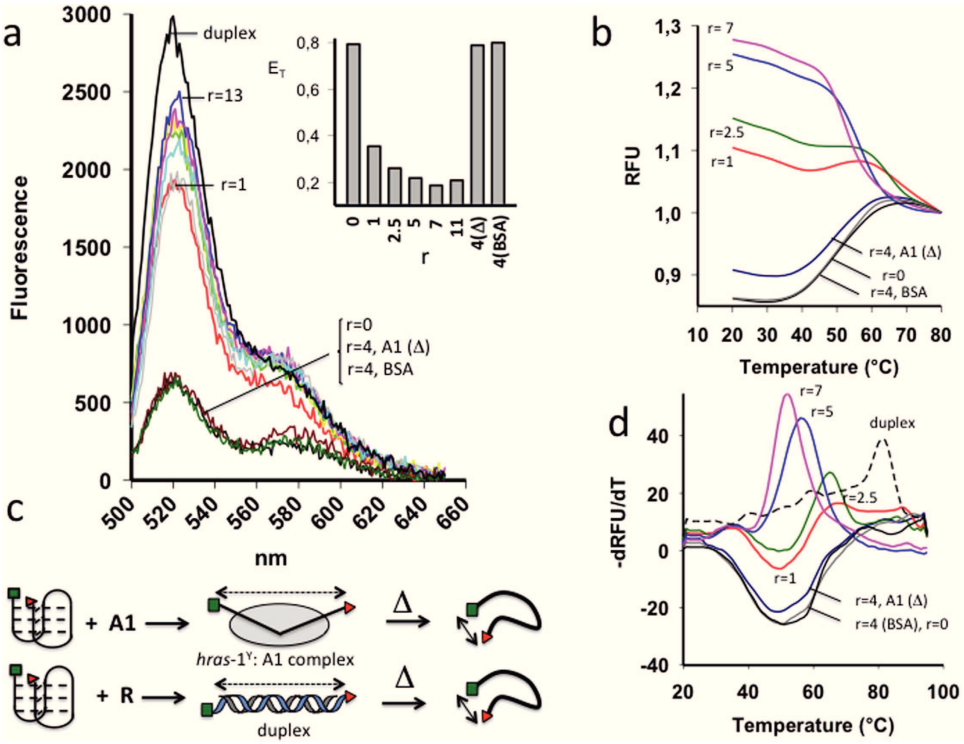



**Figure 5.** (**A**) FRET spectra of 200 nM hras-1$^{Y}$ treated with increasing amounts of purified hnRNP A1 at pH 5.5, 50 mM sodium cacodylate, 50 mM KCl. As a control BSA and denatured hnRNP A1 (A1 Δ) have been used. Note that hnRNP A1 causes a dramatic increase of the 520 nm donor emission. The emission spectra of *hras*-1$^{Y}$ hybridized to its complementary strand to yield the duplex is reported. Insight shows the energy transfer ($E_T$) between donor-acceptor as a function of hnRNP A1 concentrations; (**B**) FRET-melting of *hras*-1$^{Y}$ incubated with increasing amounts of hnRNP A1 (r = 1–7). The protein abrogates the melting profiles; (**C**) Cartoon showing the melting of *hras*-1$^{Y}$ bound to its complementary strand or to hnRNP A1 (r = 1–7); (**D**) –dRFU/dT versus T curves of *hras*-1$^{Y}$ alone, *hras*-1$^{Y}$+protein, *hras*-1$^{Y}$ in duplex, i.e. hybridized to its complementary strand.

can be seen that at pH 5.5 *hras*-1$^{Y}$ and *hras*-2$^{Y}$ (which are in the folded *i*M conformation) form with hnRNP A1 a retarded band due to a 1:1 DNA-protein complex. In the presence of 4 μg hnRNP A1, the *i*M is completely bound to the protein (Fig. 4B). With higher protein amounts, a second retarded band of much weaker intensity, probably a 1:2 complex, can be seen in the gel. When hnRNP A1 was thermally denatured before being added to the *i*M, the DNA-protein complex was abrogated and the *i*M migrated as a free molecule. As a further control, we used an unspecific protein like BSA and we found that it did not bind to the *i*M, as expected (Fig. 4C). When the *i*Ms was destabilized by replacing 4 C with 4 T (*hras*-1$^{Y}$(m)) (see Methods), the binding was significantly attenuated, suggesting that the *i*M conformation is essential for optimal hnRNP A1 binding.

**hnRNP A1 unfolds the HRAS *i*-motif.** In a previous work we have found that the *HRAS* promoter is highly active when the GC-elements are unfolded in the double-stranded conformation[3,4]. We now asked if the binding of hnRNP A1 to the *i*M involves the unfolding of this non B-DNA structure. To this purpose, we performed FRET experiments with *hras*-1$^{Y}$ tagged with ATTO-488 (donor) and TAMRA (acceptor) in 50 mM sodium cacodylate, pH 5.5, 50 mM KCl. By exciting the donor at 480 nm, both donor (520 nm) and acceptor (580 nm) emit fluorescence, as a result of FRET between the fluorophores (Fig. 5A). When *hras*-1$^{Y}$ is folded into the *i*M, the energy transfer, $E_T$, between the two fluorophores is 0.77, and their end-to-end distance is ~40 Å ($S_2$). The effect of hnRNP A1 on the *i*M was investigated by incubating protein and DNA for 1.5 h and measuring the fluorescence between 500 and 650 nm, upon donor excitation at 480 nm. It can be seen that hnRNP A1 causes a dramatic increase of the donor emission, accompanied by a decrease of $E_T$ as a function of r (r = [protein]/[*i*M]), in a dose-dependent manner, from 0.77 (r = 0) to 0.18 (r = 7) (inset). This means that the end-to-end distance in the *i*M increases from ~40 to ~64 Å, suggesting that upon binding to the protein, *hras*-1$^{Y}$ goes through a conformational change. When *hras*-1$^{Y}$ is in the duplex conformation, the fluorophores are separated by ~86 Å (26 × 0.33 = 86 Å, where 0.33 Å is the vertical rise per bp), as a 27-mer duplex behaves as an extended rigid rod. It follows that the *i*M bound to hnRNP A1 is not fully extended as in the duplex. As a control we used an unspecific protein as BSA and heated hnRNP A1 before incubation with the *i*M (20 min, 95 °C). In both cases the fluorescence of the donor did not increase, as expected (Fig. 5A).

To further support the finding that hnRNP A1 disrupts the *i*M, we carried out FRET-melting experiments, reasoning that the *i*M would not give its typical melting profile when bound to hnRNP A1. Fig. 5B shows that *hras*-1$^{Y}$ *i*M in 50 mM sodium cacodylate, pH 5.5, 50 mM KCl has a $T_M$ ~ 50 °C. In the presence of 4 equivalents (r = 4) of

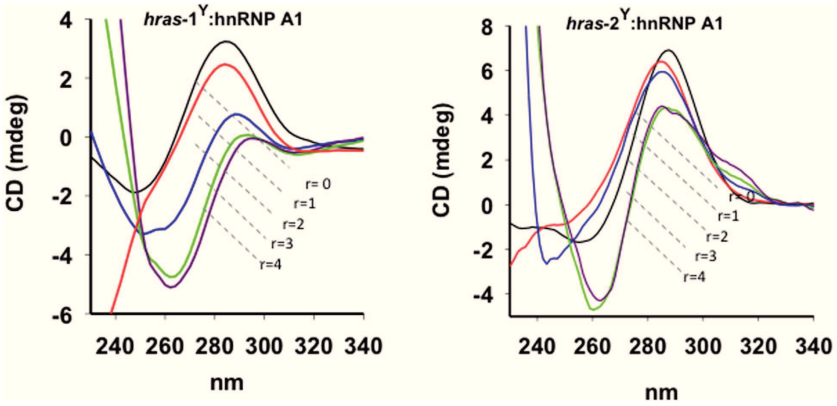



**Figure 6.** Circular dichroism analysis of 3 μM (0.5 cm pathlength cell) *hras*-$1^Y$ and *hras*-$2^Y$ at pH 5.5, 50 mM Tris-acetate, 50 mM KCl, after incubation with increasing amounts of hnRNP A1 (r = 0–4) . Spectra of DNA-protein complex have been subtracted of protein spectrum.

BSA or denatured hnRNP A1, the melting profile of *hras*-$1^Y$ did not change, as expected. In contrast, when the *hras*-$1^Y$ *i*M was incubated with increasing amounts of native hnRNP A1 (r = 1–7), its melting profile strongly changed, in keeping with the fact that *hras*-$1^Y$ bound to the protein is not in the *i*M conformation. It is worth noting that free *hras*-$1^Y$ melts with an increasing sigmoidal curve, whereas *hras*-$1^Y$ bound to hnRNP A1 melts with a decreasing sigmoidal curve (Fig. 5B). A melting profile similar to that of the *hras*-$1^Y$:hnRNP A1 complex is obtained with *hras*-$1^Y$ in the duplex conformation with its complementary strand, where the two fluorophores are separated by ~86 Å. Upon melting, the duplex releases the *hras*-$1^Y$ strand which, thanks to its flexibility, will have an end-to-end distance < 86 Å (Fig. 5C). This results in a decreasing sigmoidal melting curve, and thus in a –dRFU/dT *versus* T curve marked by a positive peak. In the same way, *hras*-$1^Y$ bound to hnRNP A1 is more extended than when it is free. Therefore, also the complex gives a decreasing melting curve and a first derivative curve with a positive peak at ~50 °C. The melting of free *hras*-$1^Y$ folded in the *i*M gives an increasing melting curve and a –dRFU/dT *versus* T curve with a negative peak (Fig. 5D). Summing up, both FRET titrations and melting provide strong evidence that hnRNP A1 unfolds the *HRAS i*M.

We have also analysed the effect of hnRNP A1 on the *hras*-$2^Y$ *i*M. We found that the protein unfolds the *i*M of *hras*-$2^Y$ at higher r values, as the *i*M formed by *hras*-$2^Y$ has a higher stability than the *hras*-$1^Y$ *i*M (58.7 *versus* 52.1 °C, at pH 5.5) ($S_3$).

The effect of hnRNP A1 on the *HRAS i*Ms was also investigated by CD (Fig. 6). At pH 5.5, *hras*-$1^Y$ and *hras*-$2^Y$ show the typical strong ellipticity at 287 nm of the *i*M conformation. When the *i*Ms are thermally denatured, the positive 287 nm ellipticity drops dramatically. This is a hallmark of the transformation of *i*M into ssDNA. A similar spectral change was obtained when we added increasing amounts of hnRNP A1 (*r* = 1, 2, 3, 4) to the *i*Ms. HnRNP A1 causes a progressive reduction of the 287 nm ellipticity, indicating that the *i*M structures are unfolded by the protein. As already observed with the FRET experiments, the unfolding effect is stronger with *hras*-$1^Y$ than with *hras*-$2^Y$, owing to the different stability of the two *i*Ms.

**Insights into the binding of hnRNP A1 to the *hras*-$1^Y$ *i*M.** Clues to the binding mode of hnRNP A1 to the *i*M can be obtained from the co-crystal structure of UP1 (the *N*-terminus of hnRNP A1 with DNA binding activity) and the human telomeric sequence $d(TTAGGG)_2$[43]. The co-crystal shows that a protein dimer binds to two single-stranded strands, in the antiparallel orientation. As the two binding domains (RRM1 and RRM2) within each protein molecule are also antiparallel, the 5′ → 3′ polarity of ssDNA with respect to the RRM orientation is the same for each RRM. The two lateral loops of the *i*M may (after a minor adjustment) provide suitable binding sites for hnRNP A1, as they are antiparallel and separated by ~15–20 Å. So, the *i*M's main function should be to provide a rigid chemical frame displaying two lateral loops with the precise nucleic acid directionality with respect to the RRM orientation. In other words, the *i*M structure should offer a kinetic advantage to the binding of hnRNP A1 (indeed, when the *i*M is disrupted, the binding is strongly attenuated, Fig. 4C). If we assume that the protein binds to the lateral loops, it can form either a 1:1 or a 1:2 complex, depending whether one or two protein molecules bind to the *i*M. The equilibria occurring in solution are: P + *i*M = P•iM (4); P + P•*i*M = $(P)_2$•*i*M (5), where P is hnRNP A1. In the presence of a large excess of P compared to *i*M (1:100), both equilibria shift to the right forming complex 1:2. In contrast, with less P (ratio 1:50) equilibrium (5) does not shift to the right and only complex 1:1 is formed (Fig. 7A, lanes 1–3). In addition, as hnRNP A1 unfolds the *i*M, its binding depends on temperature: at 0 °C only complex 1:1 is formed probably because it requires a partial unfolding of *i*M, at 37 °C complex 1:2 is favoured as it requires a complete opening of the *i*M ($S_4$). To support the binding of the protein to the lateral loops we performed the following competition experiment. As the two lateral loops of the *hras*-$1^Y$ *i*M are separated by 10 nt, complex 1:2 should be competed by an oligonucleotide containing the two lateral-loop binding sites separated by a spacer of 10 nt (with a 10 nt spacer the competitor assumes a U-shape so that two antiparallel binding sites can interact with the protein RRMs). Moreover, if the spacer of the competitor is reduced to 8, 6, 4, and 2 nt, its capacity to compete with the formation of the 1:2 complex should gradually become weaker. To test this hypothesis we designed the competitors shown in Fig. 7C. When the competitors (150-fold in excess over *i*M) were incubated with the *i*M and hnRNP A1 (100-fold over *i*M), we found that the best competitor was the oligomer containing a spacer of 10 nt,

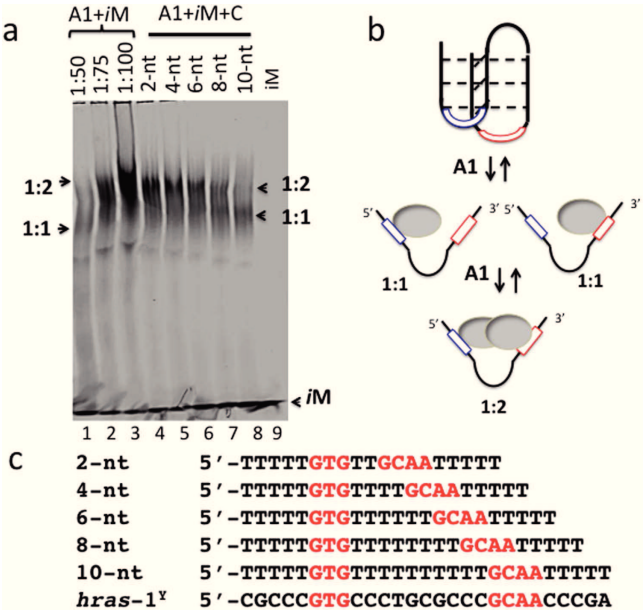



**Figure 7.** (**A**) 5% PAGE, lanes 4–8 show how the binding of hnRNP A1 to *hras*-1$^{Y}$–dy781 (Methods) (ratio 1:100) at pH 5.5 is competed by oligonucleotides 2-nt - 10-nt (150-fold over *i*M). Lanes 1–3 shows the binding of *i*M to hnRNP A1 at ratios 1:50, 1:75, 1:100. Experimental conditions: DNA and protein incubated for 40 min at 37 °C, pH 5.5 before 5% PAGE analysis; (**B**) structure of the iM with the two lateral loops to which the protein is expected to bind; (**C**) sequences of the oligonucleotide competitors 2-nt, 4-nt, 6-nt, 8-nt, 10-nt.

which corresponds exactly to the distance of the lateral-loop binding sites in the wild-type *hras*-1$^{Y}$ *i*M. These data support the binding of hnRNP A1 to the lateral loops of the *i*M, as observed for hnRNP LL and *BCL2* *i*M[6]. As stated above, the *i*M facilitates the initial binding to hnRNP A1. Then, after the *i*M is unfolded, the protein should bind more stably to the *i*M sequence.

**HnRNP A1 knockout results in downregulation of *HRAS* transcription in human bladder cancer cells.** As hnRNP A1 binds to the critical GC-elements of the *HRAS* promoter, we asked if the protein plays a role in transcription. We therefore evaluated the effect on transcription of knocking out hnRNP A1 by shRNA. First, we determined by quantitative real-time PCR the efficiency of shRNA to knock out hnRNP A1, finding that hnRNP A1 mRNA (normalized by the transcripts of β2-microglobulin and hypoxanthine guanine phosphoribosyltransferase, HPRT) was reduced to 29 ± 8% of the control (cells untreated or treated with a non-specific shRNA$^{C}$), 48 h after treatment (Fig. 8A). In the same cells we also measured the level of *HRAS* mRNA finding that when hnRNP A1 is knocked out, the *HRAS* transcription is also down-regulated to 44 ± 5% of control. This suggests that hnRNP A1 is an essential factor for transcription, as previously reported[41]. Further evidence that the proteins recognizing the *i*Ms of *HRAS* are essential for transcription was obtained with decoy oligonucleotides mimicking *hras*-1$^{Y}$ *i*M. These molecules, once introduced in the cells, should sequester the proteins (hnRNP A1 included) recognizing the C-rich strand of the *HRAS* GC-elements. To increase their nuclease resistance, the decoy oligonucleotides, namely **5291**–**5294** (see Methods), have been designed with unlocked nucleic acid (UNA) modifications (Fig. 8B) ($S_5$)[44,45]. The capacity of the UNA-modified oligonucleotides to inhibit *HRAS* transcription was investigated by quantitative real-time PCR. T24 bladder cancer cells were transfected with the decoy oligonucleotides as well as with wild-type *hras*-1$^{Y}$, using as transfecting agent jet-PEI[46]. After an incubation of 24 h, the total cellular RNA was extracted and the amount of *HRAS* mRNA relative to the housekeeping HPRT mRNA was evaluated by qRT-PCR (Fig. 8C). The results showed that oligonucleotides **5292** and **5293** reduced *HRAS* mRNA to ~50% of the control (untreated cells). We also examined by electrophoresis the nuclease resistance of the decoy oligonucleotides (Fig. 8D). The oligonucleotides were incubated in cell cultured medium containing 10% fetal bovine serum at 37 °C, pH 5.5 for 0, 18, 24 and 48 h. While *hras*-1$^{Y}$ was quickly degraded, the UNA-modified oligonucleotides, in particular **5292** and **5293,** showed a remarkable stability, as the fraction of unbroken oligonucleotide was >0.5, after 48 h of incubation. Interestingly, the enhanced activity shown by these compounds correlates nicely with their higher stability in serum.

## Conclusion

We have demonstrated that two neighbouring GC-rich elements controlling *HRAS* expression can form non B-DNA *i*M structures, which are stable under near-physiological conditions. These unusual DNA structures are recognized by hnRNP A1, one of the most abundant nuclear proteins involved in the biogenesis of RNA. We have discovered that hnRNP A1 has a clear unfolding activity against the *i*M. As the knockout of hnRNP A1 by shRNA in T24 bladder cancer cells results in the inhibition of *HRAS*, hnRNP A1 behaves as an activating transcription factor. Our data, together with those of Hurley and co-workers, who showed that hnRNP LL binds to the *i*M of

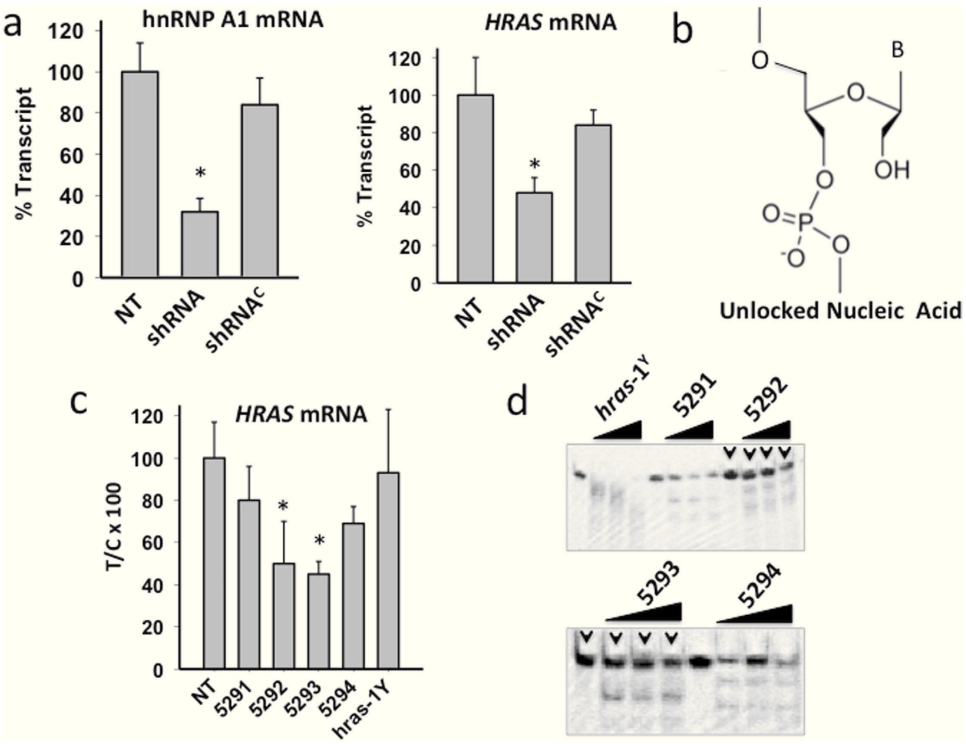



**Figure 8.** (**A**) Real-time determination of hnRNP A1 and *HRAS* mRNAs after knocking down hnRNP A1 in T24 bladder cancer cells with a specific shRNA. When hnRNP A1 is knocked down, *HRAS* mRNA is downregulated. $P < 0.05$ (*); (**B**) UNA modification introduced in the decoy oligonucleotides; (**C**) Level of *HRAS* mRNA in T24 cells treated with 200 nM *hras*-$1^Y$ or UNA-modified analogues. Total RNA was extracted 24 h after oligonucleotide transfection, retro-transcribed and subjected to real time amplification. *HRAS* mRNA expression is normalized with housekeeping gene HPRT. The percentage of residual *HRAS* mRNA compared to *HPRT* mRNA in each sample is reported. $P < 0.05$ (*); (**D**) Resistance in fetal serum of *hras*-$1^Y$ and UNA-modified analogues. Oligonucleotides have been incubated in serum for 0, 18, 24, 48 and 72 h at 37 °C. After incubation the samples have been run in denaturing PAGE, 7 M urea, 55 °C. The gels were stained with "stains all".

the *BCL2* promoter and activates transcription[6], provide the first evidence that non B-DNA *i*M structures are recognized by nuclear proteins.

The proteins of the hnRNP family have been associated with the promoter of several genes where they are supposed to participate in the transcription regulation mechanisms, although their exact role is not yet fully understood. Some of them recognize C-rich sequences in the promoters of *CMYC* (hnRNP K)[39,40], *BCL2* (hnRNP LL)[6] and *HRAS* (hnRNP A1) (present study). These proteins seem to have a complex binding capacity, as hnRNP A1 is also able to bind to G-quadruplex DNA structures in *KRAS*[17] and telomeres[47]. Remarkably, this type of binding is also associated with the disruption of G-quadruplex structures[17].

Recent mechanical folding/unfolding experiments showed that G-quadruplex and *i*M are mutually exclusive within the same double-stranded tract[48]. However, whether this also holds under *in vivo* conditions, where double stranded DNA is exposed to negative superhelicity and located in a molecular crowding environment, has not yet been demonstrated. It is possible that both G-quadruplex and *i*M are extruded from each double-stranded GC-element, in the same way as two opposing hairpins (a cruciform) are extruded from a palindromic sequence. *HRAS* could therefore be regulated by a G-quadruplex/*i*M switch that represses transcription when the structural elements are in the folded conformation. Transcription will be activated when hnRNP A1 and MAZ, which recognize the *HRAS* G-quadruplexes[3], bind to the *i*M and G-quadruplex, respectively, and then to other proteins of the transcriptional activator complex. These non B-DNA structures provide a mechanism for the control of gene expression at a different level than duplex, involving proteins recognizing these unusual structures that play a central role in gene regulation.

## Methods

**Oligonucleotides and hnRNP A1.** The oligonucleotides used in this study have been obtained from Microsynth AG (Switzerland) and Eurofins Genomics (Germany):

5′-CGCCCGTGCCCTGCGCCCGCAACCCGA (*hras*-$1^Y$)
5′-ACCGCGCGCCCCCGCCCCCGCCCCGCCCCGGCCTCG (*hras*-$2^Y$)
5′-ATTO-CGCCGCCCGTGCCCTGCGCCCGCAACCCGAGC-TAMRA (A-*hras*-$1^Y$-T)
5′-ATTO-CGC GCC CCC GCC CCC GCC CCG CCC C -TAMRA (A-*hras*-$2^Y$-T)

5′-TCG GGT TGC GGG CGC AGG GCA CGG GCG (*hras*-1$^{R}$)
5′-CGG GGC GGG GCG GGG GCG GGG GCG (*hras*-2$^{R}$)
5′- CGC TCG TGC TCT GCG CTC GCA ACT CGA (*hras*-1$^{Y}$m)
5′- TTTTTGTGTTTTTTTTTGCAATTTTT (ODN-1)
5′-CGTCCGTGTCCTGCGTCCGCAATCCGA (ODN-2)
5′-CGCCCGTGCCCTGCGCCCGCAGGGCGA (ODN-3)
5′-Dy 781-TTTTTTTCGCCCGTGCCCGTCGCCCGCAACCCGATTTTTTT-3′ (*hras*-1$^{Y}$-dy 781).

The oligonucleotides with UNA modifications have been synthesized in solid phase as previously described[44,45]: 5′-CGCCCGTGCCCTGuCGCCCGCuAACCCGuA (**5291**); 5′- CGCCCGTGCCCuUG CGCCCuGCAACCCGuA (**5292**); CGCCCGuUGCCCTGuCGCC-CGCuAACCCGuA (**5293**) and 5′-CGCCCGuUGCCCuUGCGCCCuGCAACCCGuA (**5294**) where uC, uU, uG, uA are unlocked nucleic acid nucleotides.

Recombinant hnRNP A1 tagged to GST was obtained with a high degree of purity as previously described[49] ($S_6$).

**Chromatin immunoprecipitation.** T24 urinary bladder cancer cells ($1.2 \times 10^6$) were cultured overnight in 6-cm diameter plates up to about 80% confluency and fixed in 1% formaldehyde in PBS for 5 minutes at room temperature to crosslink proteins to DNA. Chromatin immunoprecipitation assays were performed using the ChIP-IT™ Express kit (Active Motif, Rixensart, Belgium). Details are reported in $S_7$.

**ShRNA transfection, RNA extraction and real-time PCR.** T24 cells were plated in 96-well plate ($10^4$ cells/well). After 1 day we transfected the cells with hnRNP A1-specific (sc-35576-SH) and control shRNA (sc-108066) (Santa Cruz, Dallas, USA) using as transfectant agent jetPEI™ (Polyplus, NY, USA). After 48 h, RNA was extracted by using iScript ™ RT-qPCR sample preparation reagent (BioRad, USA).

For cDNA synthesis, 1.25 μl of RNA was heated at 70 °C and placed in ice. The solution was added with 7.5 μl of a mix containing (final concentrations) 1 × buffer; 0.01 M DTT (Invitrogen); 1.6 μM primer dT [MWG Biotech, Ebersberg, Germany; $d(T)_{16}$]; 1.6 μM random primers; 0.4 mM dNTPs solution containing equimolar amounts of dATP, dCTP, dGTP and dTTP (Euroclone, Pavia, Italy); 0.8 U/μl RNAse OUT; 8 U/μl of M-MLV reverse transcriptase (Life Technologies, Monza, Italy). The reactions were incubated for 1 h at 37 °C and stopped with heating at 95 °C for 5 min. As a negative control the reverse transcription reaction was performed with a sample containing DEPC water.

Real-time PCR multiplex reactions were performed with 1xKapa Probe fast qPCR kit for *HRAS* and housekeeping genes hypoxanthine-guanine phosphoribosyltransferase (HPRT) and β2-microglobulin, 2.2 μl of cDNA and primers/probes at the concentrations specified in $S_1$. The PCR cycle was: 3 min at 95 °C, 50 cycles 10 s at 95 °C, 60 s at 58 °C. Real-time PCR amplification of hnRNP A1 was performed with 1 × Kapa Sybr Fast BioRad iCycle qPCR kit (KAPA Biosystems, Wilmington, MA, USA), 300 nM of each primer, 3.5 μl of cDNA (cycle: 3 min at 95 °C, 40 cycles 10 s at 95 °C, 30 s at 58 °C). PCR reactions were carried out with a CFX-96 real-time PCR apparatus controlled by an Optical System software (version 3.1) (Bio-Rad Laboratories, CA, USA). All expressions were normalized with housekeeping genes. The sequences of the primers and probes used for the amplifications are given as supplementary data ($S_8$).

**CD and FRET experiments.** CD spectra have been obtained with a JASCO J-600 spectropolarimeter equipped with a thermostatted cell holder. CD experiments were carried out with oligonucleotides *hras*-1$^{Y}$ and *hras*-2$^{Y}$ (3 μM) in 50 mM Tris-acetate, pH from 4.5 to 8, 50 mM KCl. Spectra were recorded in 1 or 0.5 cm quartz cuvette. A thermometer inserted in the cuvette holder allowed a precise measurement of the sample temperature. The spectra have been calculated with J-700 Standard Analysis software (Japan Spectroscopic Co, Ltd) and reported as ellipticity (mdeg) *versus* wavelength (nm). Each spectrum was recorded three times, smoothed and the baseline subtracted. CD spectra of 3 μM *hras*-1$^{Y}$ and *hras*-2$^{Y}$ have been obtained also at various temperatures (20-85 °C), by both heating and cooling the sample solutions (in 50 mM sodium cacodylate pH 5, 50 mM KCl). By plotting the 287 nm ellipticity versus temperature, sigmoidal denaturing and renaturing curves were obtained, which were practically overlapping.

FRET with oligonucleotides *hras*-1$^{Y}$ and *hras*-2$^{Y}$, tagged at the 5′ and 3′ ends with ATTO-488 and TAMRA (as donor we used ATTO-488 because its pH dependence is weaker than that of FAM), were carried out on a Microplate Spectrofluorometer System (Perkin Elmer 2300 Enspire, USA). Each sample contained 50 μl dual-labelled oligonucleotide (200 nM) in 50 mM Tris-acetate buffer, pH from 4.5 to 8, 50 mM KCl and an amount of hnRNP A1 as specified in the figure captions. The samples were incubated at 37 °C as specified in the text. Emission spectra were obtained setting the excitation wavelength at 480 nm and recording the emission from 500 to 650 nm. FRET-melting experiments of *hras*-1$^{Y}$ and *hras*-2$^{Y}$ have been performed on a real-time PCR apparatus (CFX96, BioRad, Hercules, CA) in 50 mM sodium cacodylate at pH 5, 5.5, 6, 6.5 and 7, 50 mM KCl. FRET-melting experiments were obtained by increasing the temperature from 20 °C to 95 °C (0.3 °C/min). From the melting data we obtained curves reporting the fraction of folded iM against temperature. These curves were reversible (denaturing and renaturing curves overlapping). The energy transfer ($E_T$) was calculated from the fluorescence intensity of the donor D in the presence ($I_{DA}$) and absence ($I_D$) of the acceptor as:

$$E_T = 1 - \frac{I_{DA}}{I_D} \tag{6}$$

$I_{DA}$ and $I_D$ were measured in same buffer under identical concentrations ($I_D$ was obtained by transforming the dual-labeled oligonucleotide into the corresponding duplex in which the fluorophores are at a distance for which FRET = 0). The FRET efficiency values were converted to distances between donor and acceptor by using:

$$R = R_0 \sqrt[6]{\frac{1}{E_T} - 1} \tag{7}$$

where $R$ is the distance (Å) and $R_0$ is the Föster distance [defined as the distance at which energy transfer is 50% of the maximum value, assumed to be 50 Å[50]].

**Thermodynamic analysis of reversible *i*M melting curves.** The thermodynamic parameters for the folding of C-rich sequence into the *i*M conformation were obtained from the melting curves. From $\Delta G° = -RT \ln K = \Delta H° - T\Delta S°$ (8) it is obtained $\ln K = (\Delta H°/R)(1/T) + \Delta S°/R$ (9). The equilibrium constant K as a function of T is given by $K = f/(1-f)$, where $f$, the fraction of sequence folded in the *i*M conformation, is obtained form the melting curves. The plot of ln $K$ versus $1/T$ gives a straight line whose slope ($-\Delta H/R$) and y-intercept ($-\Delta S/R$) provide the thermodynamic parameters ($S_9$).

**PAGE assays.** Oligonucleotides *hras*-$1^Y$ and *hras*-$2^Y$ were end-labelled with [$\gamma$-$^{32}$P]ATP and T4 polynucleotide kinase. For competition experiments we used a DNA chemically labelled to dy-781. Before EMSA, the iM-forming oligonucleotides were allowed to form their structure in 50 mM Tris–acetate, pH 5.5, 50 mM KCl, (overnight incubation at room temperature). Radiolabelled oligonucleotides (10 nM) were incubated for 30 min at 20 °C with increasing amounts of hnRNP A1 (0–12 μg) as specified in Fig. 4B, in 50 mM Tris–acetate, pH 5.5, 50 mM KCl, 1 mM DTT, 8% glycerol, 1% Phosphatase Inhibitor Cocktail I (Sigma, Milan, Italy), 5 mM NaF, 1 mM $Na_3VO_4$, 2.5 ng/μl salmon sperm DNA (binding buffer). After incubation, the reaction mixtures were loaded in 5% PAGE in 50 mM Tris-acetate pH 5.5, thermostatted at 20 °C. After running the gel was dried and exposed to autoradiography (G E Healthcare, Milan) for 16 h at −80 °C. Mobility-shift experiments of cold *hras*-$1^Y$ and *hras*-$2^Y$ have been performed on 15% PAGE, 25 mM KCl, at pH 5 (50 mM sodium acetate) or pH 7 (50 mM Tris-acetate), 20 °C. 20% PAGE in denaturing 7 M urea conditions, was carried out in TBE. The gels were stained with "stains-all" dye. Competition assay with 28 nM *hras*-$1^Y$-dy781 were performed at 37 °C with 3 μM of hnRNP A1 (100-fold over *i*M *hras*-$1^Y$-dy781) and competitor oligonucleotides (150-fold over *i*M) in 50 mM Tris-acetate, pH 5.5, 50 mM KCl, 1 mM EDTA, 2.5ng/μl Salmon sperm. After incubation, the reaction mixtures were loaded in 5% PAGE 1xTBE, thermostated at 20 °C. After running the gel was analysed by Odyssey CLx scanner /ImageStudio Software (Li-Cor Biosciences).

**Cell culture and transfections.** T24 human urinary bladder cancer cells were maintained in exponential growth in Dulbecco's Modified Eagle's Medium (DMEM) containing 100 U/ml penicillin, 100 mg/ml streptomycin, 20 mM L-glutamine and 10% fetal bovine serum (Euroclone, Milan, Italy).

For transfection we plated 10000 cells for each well in a 96 well plate and transfected using Jet PEI (Polyplus Illkirch FRANCE) following manufacturers *in vitro* protocol for DNA oligonucleotides transfection with 400 nM oligonucleotide (48 pmol) and N/P = 3.

## References


1. Lowy, D. R. & Willumsen, B. M. Function and regulation of ras. *Annu. Rev. Biochem*. **62,** 851–891 (1993).
2. Porter, A. C. & Vaillancourt, R. R. Tyrosine kinase receptor-activated signal transduction pathways which lead to oncogenesis. *Oncogene* **17,** 13434–13452 (1998).
3. Cogoi, S., Shchekotikhin, A. E. & Xodo, L. E. HRAS is silenced by two neighboring G-quadruplexes and activated by MAZ, a zinc-finger transcription factor with DNA unfolding property. *Nucleic Acids Res.* **42,** 8379–8388 (2014).
4. Membrino, A., Cogoi, S., Pedersen, E. B. & Xodo, L. E. G4-DNA formation in the HRAS promoter and rational design of decoy oligonucleotides for cancer therapy. *PLoS One* **6,** e24421 (2011).
5. Kendrick, S., Akiyama, Y., Hecht S. M. & Hurley, L. H. The i-motif in the bcl-2 P1 promoter forms an unexpectedly stable structure with a unique 8:5:7 loop folding pattern. *J. Am. Chem. Soc*. **131,** 17667- 17676 (2009).
6. Kendrick, S. *et al.* The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Am. Chem. Soc.* **136,** 4161–4171 (2014).
7. Kang, H. J., Kendrick, S., Hecht, S. M. & Hurley, L. H. The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc*. **136,** 4172–4185 (2014).
8. Cui, Y. *et al.* Molecular population dynamics of DNA structures in a bcl-2 promoter sequence is regulated by small molecules and the transcription factor hnRNP LL. *Nucleic Acids Res.* **42,** 5755–5764 (2014).
9. Gehring, K., Leroy, J. L. & Guéron, M. A. tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* **363,** 561–565 (1993).
10. Manzini, G., Yathindra N. & Xodo, L. E. Evidence for intramolecularly folded i-DNA structures in biologically relevant CCC-repeat sequences. *Nucleic Acids Res*. **22,** 4634–4640 (1994).
11. Day, A. H., Pavlou, P. & Waller, Z. A. E. i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. & Med. Chem*. **22,** 4407–4418 (2014).
12. Mergny, J. L., Lacroix, L., Han, X., Leroy, J. L. & Hélène, C. Intramolecular folding of pyrimidine oligonucleotides into an i-DNA motif. *J. Am. Chem. Soc*. **117,** 8887–8898 (1995).
13. Mathur, V., Verma, A., Maiti, S. & Chowdhury S. Thermodynamics of i-tetraplex formation in the nuclease hypersensitive element of human c-myc promoter. *Bioch. Biophys. Res. Commun*. **320,** 1220–1227 (2004).
14. Kaushik, M., Suehl, N. & Marky, L. A. Calorimetric unfolding of the biomolecular and i-motif complexes of the human telomere complementary strand, d(C3TA2)4. *Bioph. Chemistry* **126,** 154–164 (2007).
15. Gargallo, R. Hard/soft hybrid modeling of temperature-induced unfolding processes involving G-quadruplex and i-motif nucleic acid structures. *Analytical Biochem*. **466,** 4–15 (2014).
16. Bhavsar-jog, Y. P., Dornshuld, E., van Brooks, T. A., Tschumper, G. S. & Wadkins, R. M. Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-rich DNA. *Biochemistry* **53,** 1586–1594 (2014).

17. Cogoi, S., Paramasivam, M., Spolaore, B. & Xodo, L. E. Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins. *Nucleic Acids Res*. **36,** 3765–3780 (2008).
18. Cogoi, S. & Xodo, L. E. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res*. **34,** 2536–2549 (2006).
19. Rankin, S. *et al.* Putative DNA quadruplex formation within the human c-kit oncogene. *J. Am. Chem. Soc*. **127,** 10584–10589 (2005).
20. Fernando, H. *et al.* A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. *Biochemistry* **45,** 7854–7860 (2006).
21. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. & Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. USA* **99,** 11593–11598 (2002).
22. Hurley, L. H., Von Hoff, D. D., Siddiqui-Jain, A. & Yang, D. Drug targeting of the c-MYC promoter to repress gene expression via a G-quadruplex silencer element. *Semin Oncol*. **33,** 498–512 (2006).
23. Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug. Discov*. **10,** 261–275 (2011).
24. Brooks, T. A. & Hurley, L. H. The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat. Rev. Cancer* **9,** 849–861 (2009).
25. Bucek, P., Jaumot, J., Aviñó, A., Eritja, R. & Gargallo, R. pH-Modulated Watson–Crick Duplex–Quadruplex Equilibria of Guanine-Rich and Cytosine-Rich DNA Sequences 140 Base Pairs Upstream of the *c-kit* Transcription Initiation Site. *Chemistry* **15,** 12663–12671 (2009).
26. Xu, Y. & Sugiyama, H. Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb). *Nucleic Acids Res*. **34,** 949–954 (2006).
27. Guo, K. *et al.* Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene. *J. Am. Chem. Soc*. **129,** 10220–10228 (2007).
28. Sun, D. & Hurley, L. H. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J. Med. Chem*. **52,** 2863–74 (2009).
29. Selvam, S., Koirala, D., Yu, Z. & Mao, H. Quantification of topological coupling between DNA superhelicity and G-quadruplex formation. *J. Am. Chem. Soc*. **136,** 13967–13970 (2014).
30. Miyoshi, D., Matsumura, S., Nakano, S. & Sugimoto, N. Duplex dissociation of telomere DNAs induced by molecular crowding. *J. Am. Chem. Soc*. **126,** 165–169 (2004).
31. Zhao, C., Ren, J. & Qu, X. Single-walled carbon nanotubes binding to human telomeric i-motif DNA under molecular-crowding conditions: more water molecules released. *Chemistry* **14,** 5435–5439 (2008).
32. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).
33. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324,** 1029–1033 (2009).
34. Saenger, W. *Principles of Nucleic Acid Structures*. (Ed. Springer-Verlag, New York 1984).
35. Chen, C. Study of pH-induced folding and unfolding kinetics of the DNA i-motif by stopped-flow circular dichroism. *Langmuir* **28,** 17743–17748 (2012).
36. Dhakal, S. *et al.* Coexistence of an ILPR i-motif and a partially folded structure with comparable mechanical stability revealed at the single-molecule level. *J. Am. Chem. Soc*. **132,** 8991–8997 (2010).
37. Smiatek, J., Chen, C., Liu, D. & Heuer, A. J. Stable conformations of a single stranded deprotonated DNA i-motif. *J. Phys. Chem. B*. **115,** 13788–13795 (2011).
38. Yang, B. & Rogers, M. T. Base-pairing energies of proton-bound heretodimers of cytosine and modified cytosines: implications for the stability of DNA i-motif conformations. *J. Am. Chem. Soc*. **136,** 282–288 (2014).
39. Michelotti, E. F., Michelotti, G. A., Aronsohn, A. I. & Levens, D. Heterogeneous nuclear ribonucleoprotein K is a transcription factor. *Mol Cell Biol*. **16,** 2350–2360 (1996).
40. Takimoto, M. *et al.* Specific binding of heterogeneous ribonucleoprotein particle protein K to the human c-myc promoter, *in vitro*. *J. Biol. Chem*. **268,** 18249–18258 (1993).
41. Jean-Philippe, J., Paz, S. & Caputi, M. hnRNP A1: the Swiss army knife of gene expression. *Int. J. Mol. Sci*. **14,** 18999–19024 (2013).
42. Ushigome, M. *et al.* Up-regulation of hnRNP A1 gene in sporadic human colorectal cancers. *Int. J. Oncol*. **26,** 635–640 (2005).
43. Ding, J. *et al.* Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev*. **13,** 1102–1115 (1999).
44. Campbell, M. A. & Wengel, J. Locked vs. unlocked nucleic acids (LNA vs. **UNA**): contrasting structures work towards common therapeutic goals. *Chem. Soc. Rev*. **40,** 5680–5689 (2011).
45. Pasternak, A. & Wengel, J. Unlocked nucleic acid–an RNA modification with broad potential. *Org. Biomol. Chem*. **9,** 3591–3597 (2011).
46. Brunner, S., Fürtbauer, E., Sauer, T., Kursa, M. & Wagner, E. Overcoming the nuclear barrier: cell cycle independent nonviral gene transfer with linear PEI or electroporation. *Mol. Ther*. **5,** 80–86 (2002).
47. Fukuda, H. *et al.* Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1. *Proc. Natl. Acad. Sci. USA* **99,** 12685–12690 (2002).
48. Dhakal, S. *et al.* G-quadruplex and i-motif are mutually exclusive in ILPR double-stranded DNA. *Biophys J*. **102,** 2575–2584 (2012).
49. Paramasivam, M. *et al.* Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: implications for transcription. *Nucleic Acids Res*. **37,** 2841–2853 (2009).
50. Clegg, R. M. Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol*. **211,** 353–388 (1990).


## Acknowledgements

This work was supported by AIRC (The Italian Association for Cancer Research; IG2013, Project Code 14301). We thank Dolores Ross for proofreading the manuscript.

## Author Contributions

L.X. and S.C. conceived the study, G.M. and S.C. performed the experiments, E.P. designed and synthesized the decoys, L.X. wrote the main manuscript text, all authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at http://www.nature.com/srep

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article**: Miglietta, G. *et al.* GC-elements controlling *HRAS* transcription form *i*-motif structures unfolded by heterogeneous ribonucleoprotein particle A1. *Sci. Rep.* **5**, 18097; doi: 10.1038/srep18097 (2015).


This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

# Supporting Information

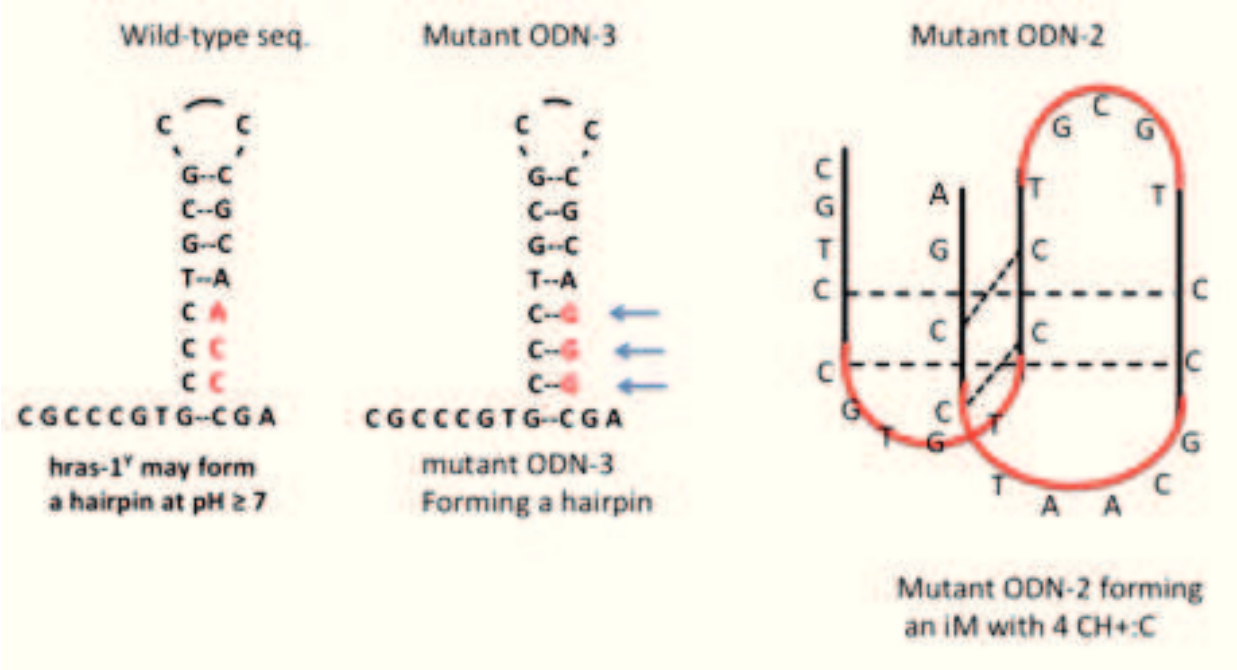
GC-rich elements controlling *HRAS* transcription form *i*-motif structures unfolded by heterogeneous ribonucleoprotein particle A1

Giulia Miglietta,[a,&] Susanna Cogoi,[a,&] Erik B. Pedersen[b] and Luigi E. Xodo [a*]

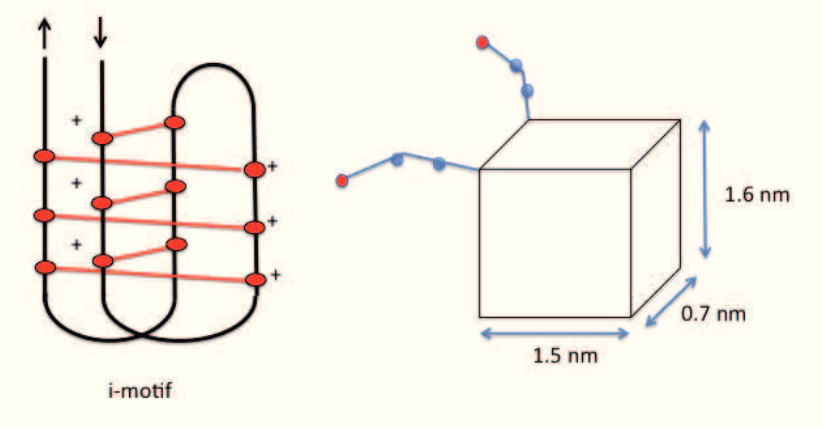
[a] Department of Medical and Biological Sciences, P.le Kolbe 4, 33100 Udine, Italy;

[b] Nucleic Acid Center, Institute of Physics and Chemistry, University of Southern Denmark, DK-5230 Odense M, Denmark.

**Supporting data S1**: Sequence of wild type *hras*-1$^{Y}$ that form an *i*M in the pH range between 5 and 7. At pH ≥ 7 the sequence may form a hairpin stabilized by a stem with 5 W.C. base pairs. Sequence of mutant ODN-3 forming a hairpin stabilized by a stem of 7 base pairs. Sequence and possible structure of mutant ODN-2 forming an *i*M stabilized by 4 CH$^{+}$:C base pairs.



**Supporting data S2**: The inter-phosphate distance across the wide and narrow groove are respectively 1.5 and 0.7 nm (1-3). The high of the *i*M was estimated by assuming $CH^+$:C stacking interval of 0.32 nm (4). The distance between two consecutive nt is assumed to be 0.43 nm (5). The end-to-end distance should be 7+4x4.3=24 Å, without taking into account the dyes. Considering that each dyes is conjugated at the 5' and 3' ends through a spacer, the end-to-end distance given by FRET, 40 Å, is consistent with the size of the *i*M.





1. Mills, M.; Lacroix, L.; Arimondo, P. B.; Leroy, J. L.; Francois, J. C.; Klump, H.; Mergny, J. L., *Curr. Med. Chem. - Anti-Cancer Agents* 2002, 2, 627-644.

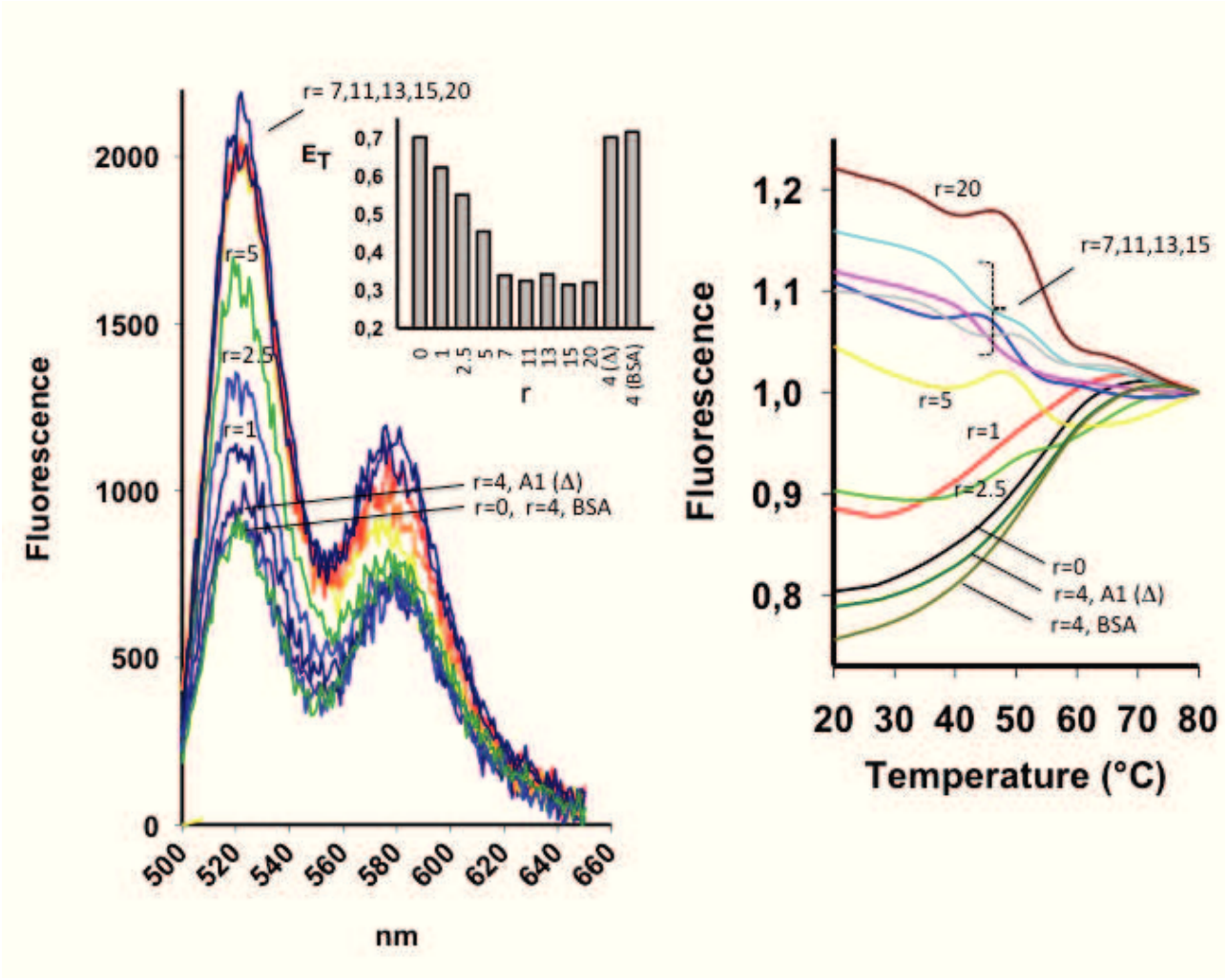
2. Kang, C.; Berger, I.; Lockshin, C.; Ratliff, R.; Myozis, R.; Rich, A., *Proc. Nat. Acad. Sci. USA* 1995, 92, 3874-3878.

3. Cai, L.; Chen, L.; Raghavan, S.; Ratliff, R.; Myozis, R.; Rich, A., *Nucleic Acids Res.* 1998, 26, 4696-4705.

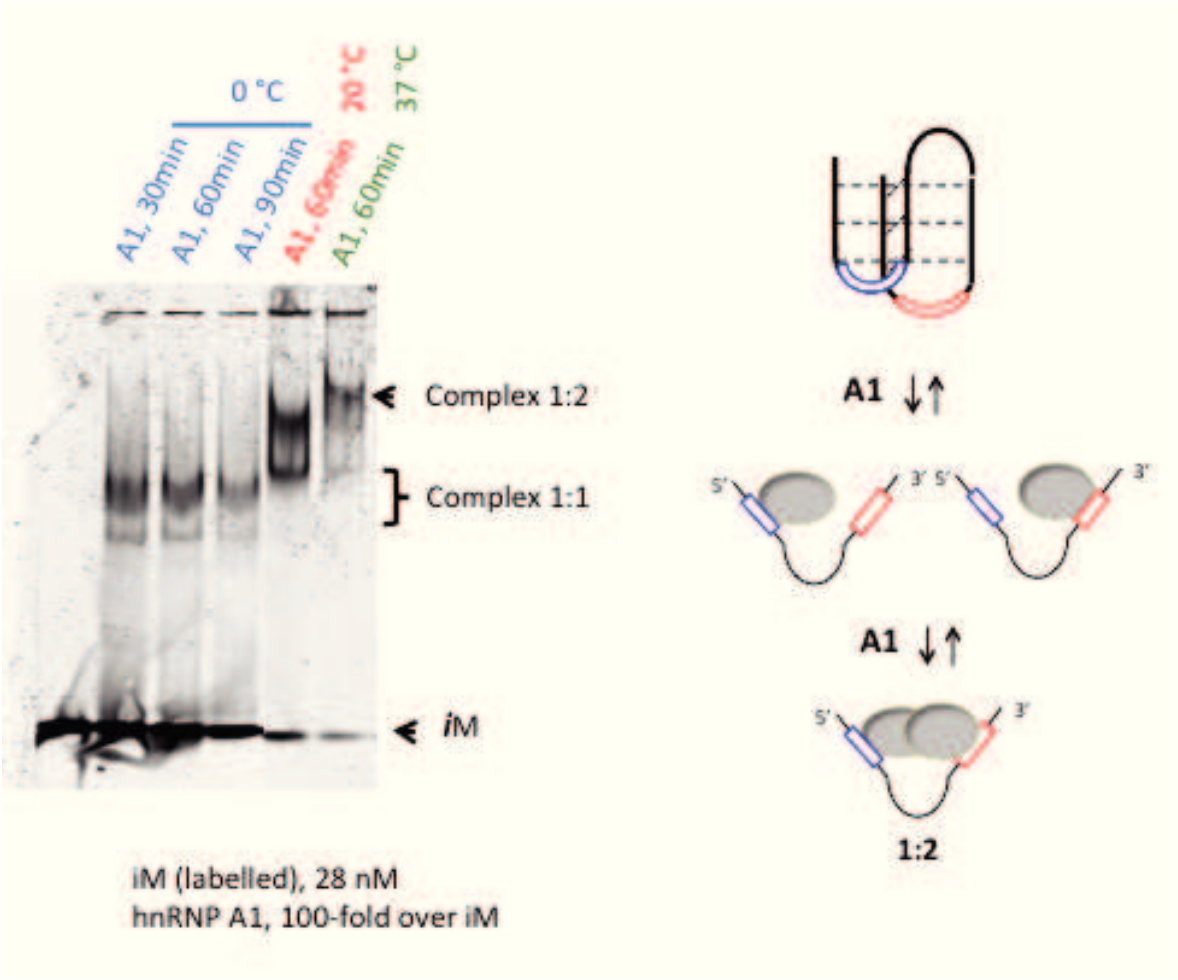
4. Esmaili, N.; Leroy, J. L., *Nucleic Acids Res.* 2005, 33, 213-224.

5. Dhakal, S., Schonhoft, J.D., Koirala, D., Yu, Z., Basu, S. and Mao, H. (2010) *J. Am. Chem. Soc.* 132, 8991–8997.

**Supporting data S3:** (Left) FRET spectra of 200 nM *hras*-$2^{Y}$ treated with increasing amounts of purified hnRNP A1 at pH 6.5, 50 mM sodium cacodylate, 50 mM KCl. As a control BSA and denatured hnRNP A1 (A1 Δ) have been used. HnRNP A1 causes a dramatic increase of the 520 nm donor emission. Insight shows the energy transfer ($E_T$) between donor-acceptor as a function of hnRNP A1 concentrations; (B) FRET-melting of *hras*-$2^{Y}$ incubated with increasing amounts of hnRNP A1. The protein abrogates the melting profiles.



**Supporting data S4.** Binding of hnRNP A1 (100-fold) to the *i*M (28 nM) in binding buffer at pH 5.5 and T= 0, 20 and 37 °C. The iM sequence is 5'-Dy 781-TTT TTT TCG CCC GTG CCC GTC GCC CGC AAC CCG ATT TTT TT-3' (*hras*-1$^{Y}$-dy781). Before incubation with hnRNP A1 it was allowed to form its structure in 50 mM Tris-acetate pH 5.5 and 50 mM KCl. Then *hras*-1$^{Y}$-dy781was incubated for 40 min at 0, 20 or 37 °C with 3 µM of hnRNP A1 (100-fold) in 50 mM Tris-acetate, pH 5.5, 50 mM KCl, 1 mM EDTA, 2.5ng/µl Salmon sperm. After incubation, the reaction mixtures were loaded in 5% PAGE 1 x TBE thermostated at 20°C. After running the gel was analysed by Odyssey CLx scanner and the shapes quantified by ImageStudio Software (Li-Cor Biosciences).



**Supporting data S5:** UNA decoy oligonucleotides**.**

The decoy oligonucleotides used are:

5’-CGCCCGTGCCCTGuCGCCCGCuAACCCGuA (**5291**);

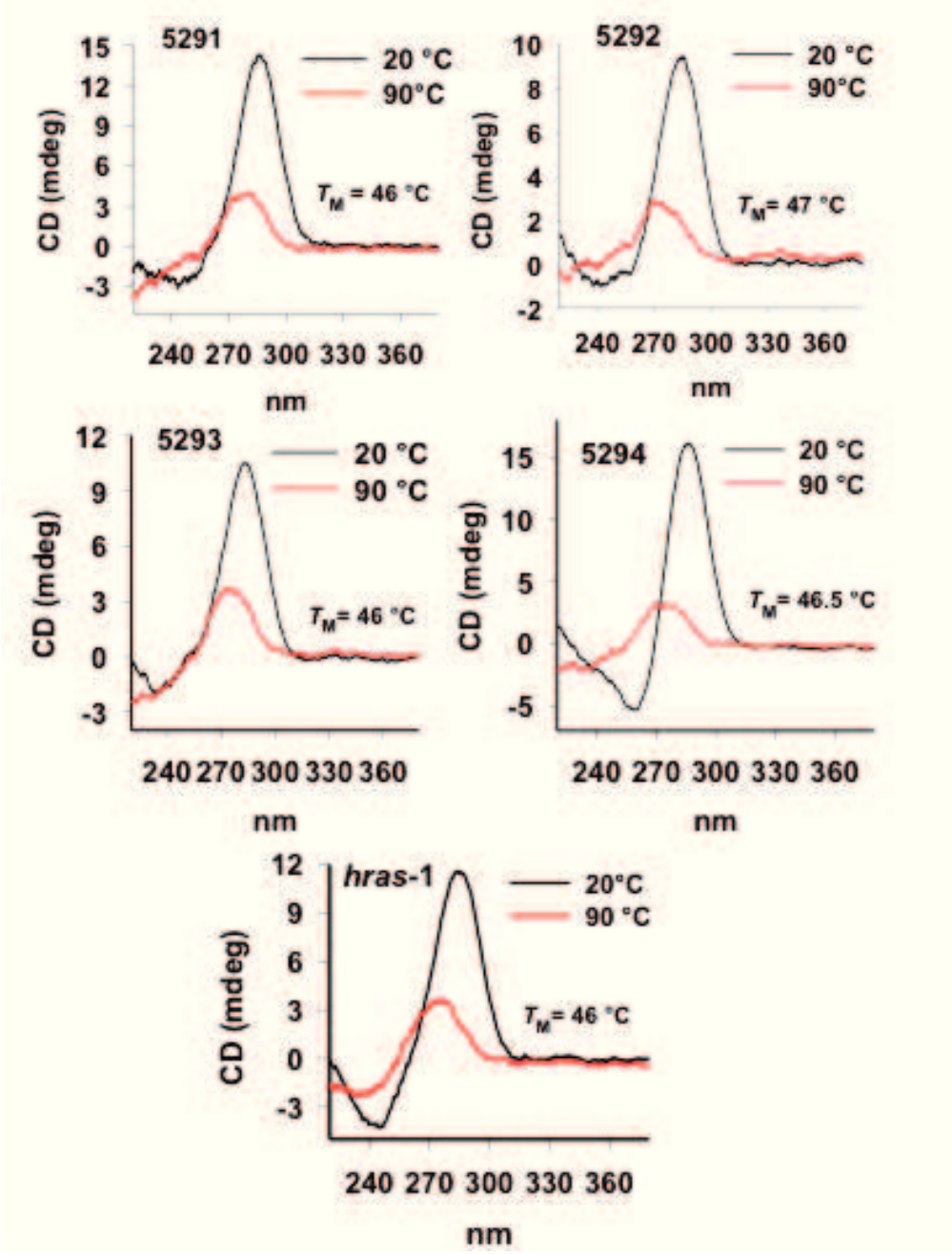
5’- CGCCCGTGCCCuUGCGCCCuGCAACCCGuA (**5292**);

5’-CGCCCGuUGCCCTGuCGCCCGCuAACCCGuA (**5293**)

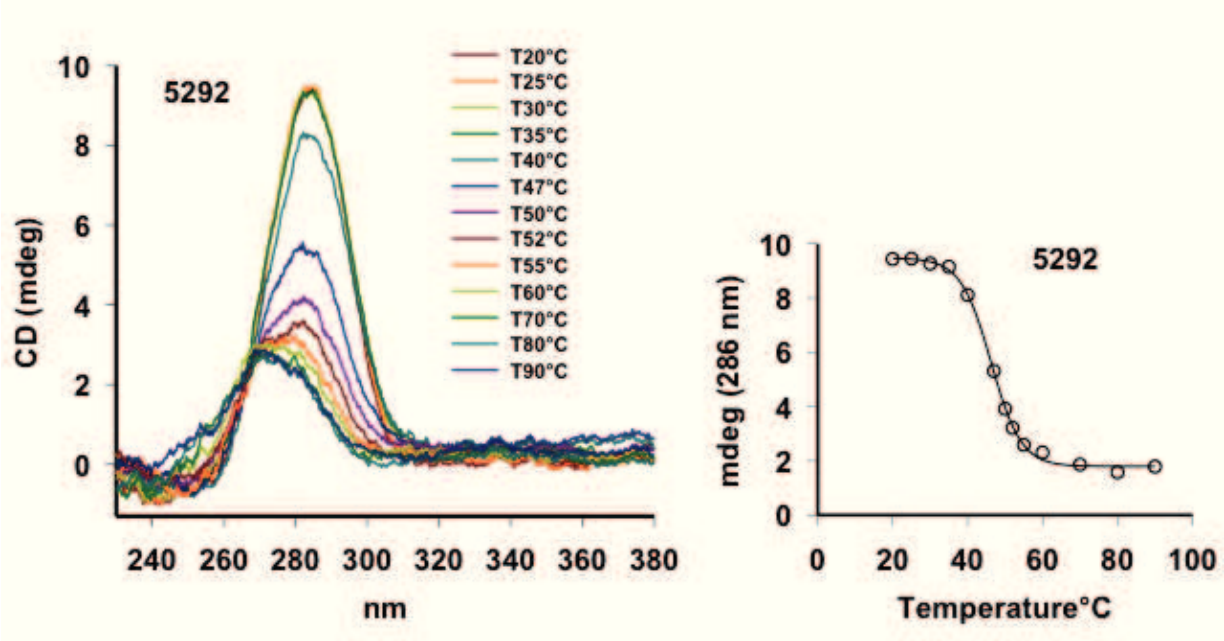
5’-CGCCCGuUGCCCuUGCGCCCuGCAACCCGuA (**5294**)

where uC, uU, uG, uA are unlocked nucleic acid nucleotides.

**A**



**B**




**(A)** Spectra of UNA-modified and wild-type *hras*-1$^{Y}$ oligonucleotides (3 μM) at 20 and 90 °C in 50 mM cacodylate pH 5.5, 50 mM KCl. Note that the UNA modifications do not alter the structure of the *i*M. For each sequence we report the melting temperature obtained by CD spectra measured as a function of temperature. In panel B we show a typical CD-melting experiment obtained with oligonucleotide 5292.

**Supporting data S$_6$: Recombinant hnRNP A1**

In brief, hnRNP A1 tagged to GST was expressed in *Escherichia coli* BL21 using pGEX-hnRNP A1. After transformation, the bacteria were grown at 37°C in the presence of 50 μg/ml ampicillin up to an absorbance at 600 nm of 0.8 before the IPTG treatment (100 μM final concentration). The cells were allowed to grow for 5 h, harvested and centrifuged at 5000 r.p.m., 4°C. The supernatant was removed carefully and the pellet was re-suspended in a solution of PBS with 20 μM PMSF and 5 mM DTT. The bacteria were lysed by sonication (3 times 30 s, 20 kHz, Bioruptor, Diagenode, NY), added with Triton X-100 (1% final concentration) and incubated for 1.5 h on a shaker at room temperature. The lysate was then centrifuged for 10 min at 4°C at 10 000 r.p.m. Glutathione Sepharose 4B (GE Healthcare) (50% slurry in PBS) was added to the supernatant and incubated for 1.5 h at 4°C on a shaker. The mix was centrifuged for 5 min at 500 g and the pellet was washes 3 times in PBS and eluted with elution buffer containing 20 mM NaCl, 20 mM reduced glutathione, 200 mM Tris–HCl, pH 9.5. Finally, purification of hnRNP A1 was checked by SDS–PAGE.

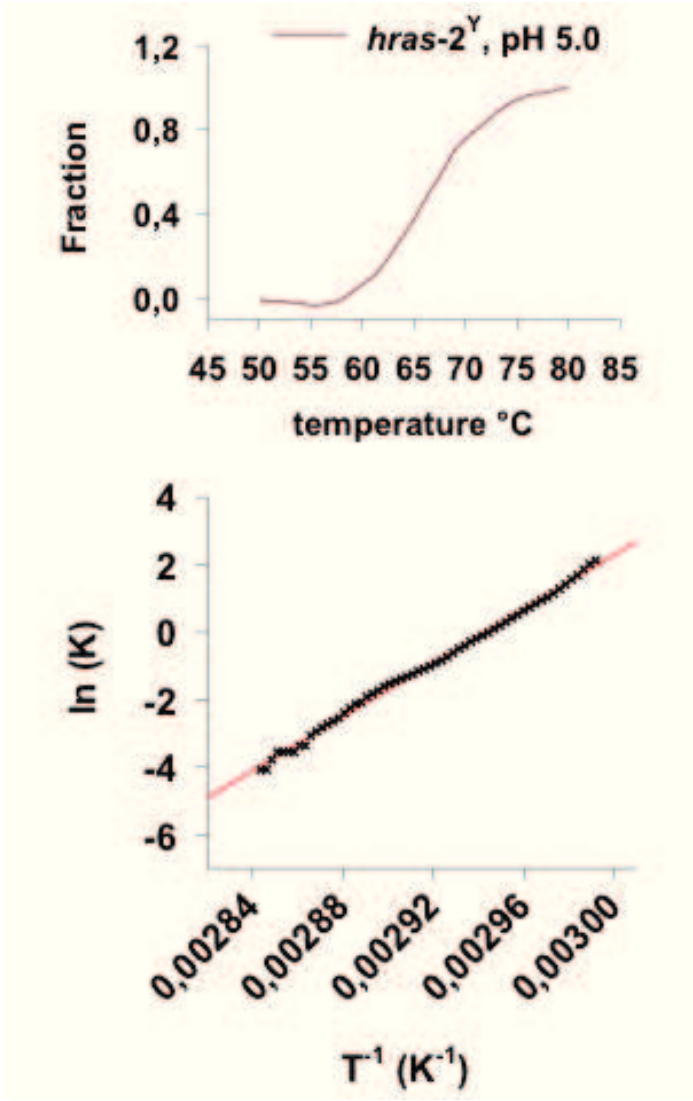
**Supporting data S7: Chromatin Immunoprecipitaion**

T24 urinary bladder cancer cells (1.2 x $10^6$) were cultured overnight in 6-cm diameter plates up to about 80% confluency and fixed in 1% formaldehyde in PBS for 5 minutes at room temperature to crosslink proteins to DNA. Chromatin immunoprecipitation assays were performed using the ChIP-IT™ Express kit (Active Motif, Rixensart, Belgium). The cells were lysed in 100 mM NaCl, 20 mM Tris-HCl, pH 7.4, 1 mM EDTA, 0.5 % NP-40, 0.5 % Na-deoxycholate, 0.1 SDS and sonicated to fragment chromosomal DNA into ~ 500 bp (8 min, 20 kHz, Bioruptor, Diagenode, NY). Sheared chromatin (6 μg in each sample) was diluted and incubated overnight, at 4 °C, with 0.5 μg of each specific antibody (Ab) and recovered with Protein G magnetic beads. The Abs used are hnRNP A1, (Sigma Aldrich), negative control mouse IgG (Active Motif), positive control RNA Pol II mouse monoclonal Ab (Active Motif). After incubation, the mixtures were spinned and the beads washed 3-times with kit washing buffers. The chromatin was treated with kit elution buffer for 15 min at RT on a shaker. We then used proteinase K (37 °C) to reverse the crosslinking and after 1 h we stopped the reaction with a proteinase K stop solution. The DNA recovered was amplified by PCR, using the following primers (accession number J00277): (i) 5'-GGCTCCTGACAGACGGG (304-320; *hras*-1for) and 5'-GCATGGGCTCCGTCC (477-491; *hras*-1rev) giving a 188 bp product; (ii) 5'-GGACGGAGCCCATGC (477-491; *hras*-2for) and 5'-CGTATTGCTGCCGCCT (624-639; *hras*-2rev) giving a 163 bp product. Amplification of control G-sequence (133 bp) downstream of the GC-elements was performed with: 5'-TCATTAAGAGCAAGTGGGGGC (1407-1428; control-for), 5'-CATCTGAAGGGCAAACCCACA (1518-1539, control-rev). Amplification products were separated by 8 % acrylamide gel in TBE and quantified with a Gel–DOC apparatus (Bio-Rad Laboratories, CA, USA). The data have been normalized by IgG.

**Supporting data S8**: Sequences of primers and probes used in this study

| Accession Number | SENSE | ANTI-SENSE | PROBE |
|---|---|---|---|
| NM_001130442<br>(HRAS) | **GCTGATCCAGAACCATTT**<br>(from 254 to 271) | **GTATCCAGGATGTCC AAC**<br>(from 344 to 361) | **HEX-CAATGACCACCTGCTTCCG-BHQ2**<br>(from 309 to 327) |
| NM_004048<br>(β2-microglobulin) | **CCCCACTGAAAAAGATGA**<br>(from 333 to 350) | **CCATGATGCTGCTTAC AT**<br>(from 415 to 432) | **ROX-TATGCCTGCCGTGTGAACC-BHQ2**<br>(from 352 to 370) |
| NM_000194<br>(HPRT) | **CTTGATTGTGGAAGATATAA TTG**<br>(from 557 to 575) | **TATATCCAACACTTCG TGG**<br>(from 672 to 690) | **Cy5-CTTGCGACCTTGACCATCTT-BHQ2**<br>(from 633 to 652) |
| NM_002136<br>(hnRNP A1) | **CATCGTTAAAGTCTCTCT TCAC**<br>(from 74 to 95) | **CAGGCTCTCATCAGTT GT**<br>(from 177 to 194) | |

**Supporting data S9**: (Top panel) Denaturation profile of *hras*-$2^{Y}$ in 50 mM sodium cacodylate pH 5, 50 mM KCl, obtained from FRET-melting experiments. The profile is reversible; (B) Typical thermodynamic analysis of the denaturation profile according to a two-state model. The parameters obtained are: ΔH= -338.2 kJ/mol, ΔS= -0.9 kJ/mol K, ΔG= -29.6 kJ/mol.



**Section B:**

# Nucleic acid targeted therapy: G4 oligonucleotides downregulate HRAS in bladder cancer cells through a decoy mechanism

G. Miglietta, AS. Gouda, S. Cogoi, EB. Pedersen, LE. Xodo

ACS medicinal chemistry letters 6 (12), 1179-1183

ACS Medicinal Chemistry Letters

Letter

pubs.acs.org/acsmedchemlett

# Nucleic Acid Targeted Therapy: G4 Oligonucleotides Downregulate *HRAS* in Bladder Cancer Cells through a Decoy Mechanism

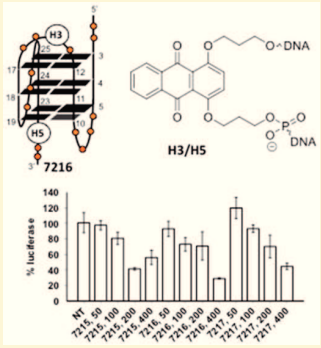
Giulia Miglietta,[†,§] Alaa S. Gouda,[‡,§,∥] Susanna Cogoi,[†] Erik B. Pedersen,[‡] and Luigi E. Xodo*[,†]

[†]Department of Medical and Biological Sciences, University of Udine, 33100 Udine, Italy

[‡]Nucleic Acid Center, Institute of Physics and Chemistry, University of Southern Denmark, DK-5230 Odense M, Denmark

Supporting Information

**ABSTRACT:** In a previous study we have demonstrated that two neighboring G-quadruplexes, *hras*-1 and *hras*-2, located immediately upstream of the major transcription start site of *HRAS*, bind MAZ, a nuclear factor that activates transcription (Cogoi, S.; et al. *Nucl. Acid Res.* **2014**, *42*, 8379). For the present study we have designed G4 oligonucleotides with anthraquinone insertions and locked nucleic acids (LNA) modifications mimicking quadruplex *hras*-1. Luciferase, qRT-PCR, and Western blot data demonstrate that these constructs efficiently down regulate *HRAS* in T24 bladder cancer cells. The inhibitory efficiency of the G4 oligonucleotides correlates with their nuclease resistance in the cell environment. By chromatin immunoprecipitation we show that the association of MAZ to the *HRAS* promoter is strongly attenuated by the designed G4 oligonucleotides, thus suggesting that these constructs behave with a decoy mechanism.




**KEYWORDS:** *HRAS, G4-oligonucleotides, anthraquinone insertions, T24 bladder cancer cells, decoy mechanism*

Mutations in the ras genes have been associated with about one-third of all human cancers.[1] In particular, mutations in codon 12 of *HRAS* are found in more than 35% of bladder cancers.[2,3] The ras genes encode for a GTP-binding protein activating downstream signal pathways that control several cell functions including proliferation.[3] The involvement of mutant *HRAS* in the pathogenesis of bladder cancer is well established.[4−6] Bladder cancer is commonly treated with cisplatin-based combination therapies, which, however, develop drug resistance. As this limits the efficacy of the therapy,[7] innovative therapeutic strategies are desirable. Given the central role of mutant *HRAS* in the pathogenesis of bladder cancer, the downregulation of this oncogene should inhibit cell growth and/or sensitize cancer cells to chemotherapy.[8] To address this issue, we focused on the *HRAS* region immediately upstream of the major transcription start site (TSS), including two neighboring G-rich quadruplex-forming sequences, namely, *hras*-1 and *hras*-2, which play a critical role in transcription regulation (Figure 1).[9,10] These two sequences overlap binding sites for MAZ and Sp1, two zinc-finger transcription factors that are essential for *HRAS* expression. By FRET experiments we demonstrated that MAZ upon binding to the promoter unfolds the G-quadruplexes.[10] We proposed a regulatory mechanism according to which the two quadruplex-forming sequences behave as a molecular switch for the control of gene expression. Under normal conditions, *HRAS* transcription is blocked by the two G-quadruplexes. Transcription is then activated when MAZ binds to and unfolds the G-quadruplexes, thus favoring the formation of the activated transcription complex (Figure 1).[10,11] In the light of this regulatory model we have

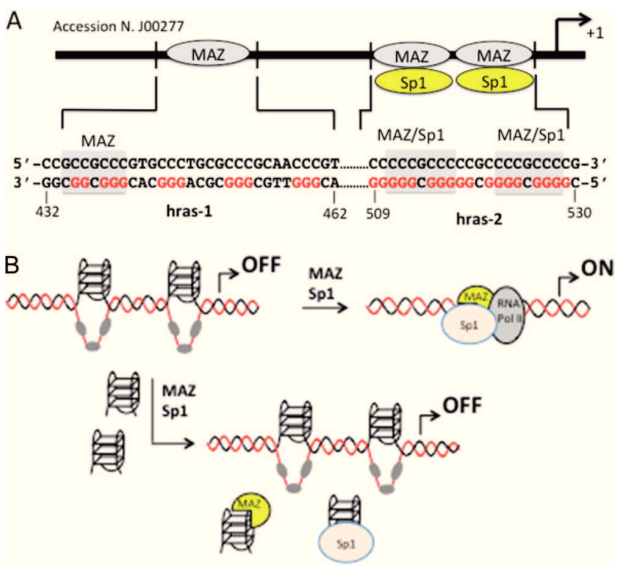



**Figure 1.** (A) Sequence of the *HRAS* promoter upstream of TSS. Two neighboring G4 sequences, *hras*-1 and *hras*-2, form a molecular switch that control transcription. (B) Decoy strategy.

hypothesized a decoy strategy to inhibit oncogenic *HRAS* in human bladder cancer cells. According to this approach, when

**Received:** August 4, 2015
**Accepted:** October 18, 2015
**Published:** October 18, 2015

ACS Publications © 2015 American Chemical Society



DOI: 10.1021/acsmedchemlett.5b00315
*ACS Med. Chem. Lett.* 2015, 6, 1179−1183

**Table 1. Sequences of the Designed Anthraquinone G4 Decoys**

| G4-decoy | 5' → 3' [a] | Mw calculated | Mw [b] Measured | ε (260 nm) | $T_M$ [c] (°C) | Topology [d] |
|---|---|---|---|---|---|---|
| 7213 | TCGGGTTHCGGGCGCAGGGCHCGGGCGG | 9044.78 | 9045.74 | 268.72 | 63.8 | Antiparallel |
| 7214 | THGGGTTGCGGGCGCAGGGCACGGGHGG | 9108.78 | 9109.64 | 278.62 | 56.8 | mixed P/A |
| 7215 | THGGGTTGCGGGCHCAGGGCACGGGCGG | 9040.82 | 9040.68 | 275.12 | 60.0 | Antiparallel |
| 7216 | TCGGGTTGCGGGCHCAGGGCACGGGHGG | 9040.82 | 9042.63 | 275.12 | 55.0 | mixed P/A |
| 7217 | TCGGGTTGCGGGCGCAGGGCACGGGCGG | 8852.52 | 8853.18 | 276.50 | 56.2 | Antiparallel |

[a]H = 1,4-anthaquinone insertion; underlined bases are LNA modified. [b]Molecular weight measured by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF). [c]$T_M$s from UV-melting profiles of compounds in 100 mM KCl and 20 mM phosphate buffer, pH 7.4. [d]Topology of the G4 decoys determined by circular dichroism.

short DNA fragments mimicking one of the *HRAS* G-quadruplexes are delivered to the cells, they should compete with the binding of the transcription factors MAZ and Sp1 to the promoter, with the result that transcription will be inhibited.

This nucleic-acid target strategy was first applied to block the binding of NF-kB and STAT3 to the corresponding *cis*-elements.[12−17] Recently, we employed decoy oligonucleotides against *KRAS* in Panc-1 cells obtaining very promising results both in vitro and in vivo.[18] To suppress *HRAS* in cancer cells by a decoy strategy, we designed quadruplex-forming (G4) oligonucleotides with locked nucleic acids (LNA) modifications and anthraquinone insertions.

The primary structure of these compounds is reported in Table 1. The LNA modifications should increase the resistance to nucleases,[19] whereas the anthraquinone insertions, by promoting $\pi$-stacking interactions, should increase the stability of the folded decoy oligonucleotides. The anthraquinone chromophore is a well-known intercalator that has been used in the development of G4 conjugates.[20] To insert the anthraquinone moiety into specific positions of the decoy oligonucleotides, we used a 1,4-substituted anthraquinone phosphoramidite[21] and synthesized the decoy oligonucleotides in solid phase (Supporting Information, S1). The anthraquinone-modified G4 oligonucleotides have been purified by HPLC and their molecular weights verified by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry analysis (Table 1). The calculated and measured molecular weights were in excellent agreement, with a difference of <0.03%. The G4 decoys have been designed with the sequence of the promoter *HRAS* segment called *hras*-1 (Figure 1), which is known to assume an antiparallel G-quadruplex in the presence of KCl (S2).[9,10] All the compounds were engineered with two or three LNA modifications and two anthraquinone insertions, each replacing one nucleotide, in positions H2/H4 (**7213**), H1/H5 (**7214**), H1/H3 (**7215**), and H3/H5 (**7216**) (Figure 2A,B). Compound **7217** was instead designed with only LNA modifications. By CD we examined if the insertions modified the conformation of the G-quadruplex formed by *hras*-1. Figure 2C shows the CD spectra of the designed compounds in 100 mM KCl, 20 mM phosphate buffer, pH 7.4. It can be seen that **7213**, **7215**, and **7217** maintain the antiparallel conformation with positive and negative ellipticities at 290 and 260 nm,[17] whereas compounds **7214** and **7216** adopt a mixed parallel/antiparallel (P/A) form characterized by two positive ellipticities at 290 and 260 nm.[22] Thus, the positions of the anthraquinone insertions influence

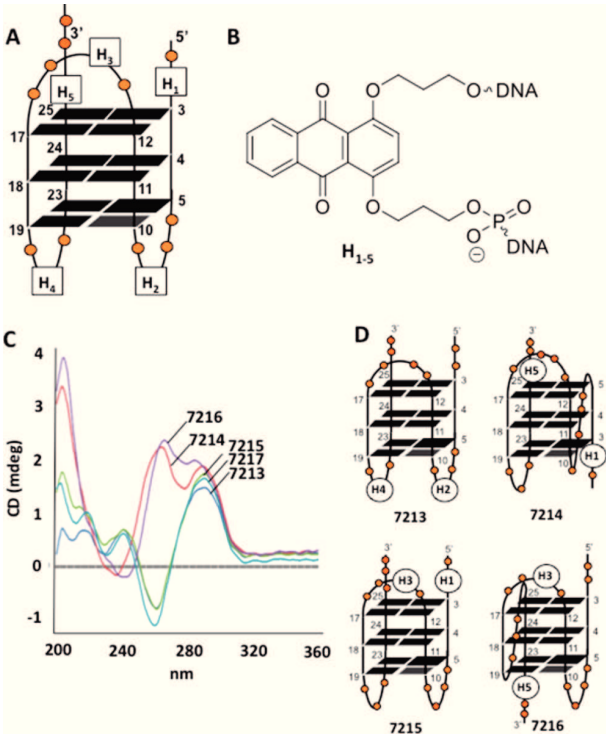



**Figure 2.** (A) Putative structure of the designed decoy oligonucleotides reported in Table 1 with LNA modifications and anthraquinone insertions H2/H4 (**7213**), H1/H5 (**7214**), H1/H3 (**7214**), H3/H5 /(**7216**). (B) Structure of the anthraquinone insertion.[21] (C) CD spectra of the designed decoy oligonucleotides in 20 mM phosphate buffer, pH 7.4, 100 mM KCl. (D) putative structure of the decoy oligonucleotides, H1−H5 = anthraquinone insertions.

the folding of the resulting G4 oligonucleotides. The change in G4 topology modified the electrophoretic mobility of the compounds. Figure 3A shows that the two oligonucleotides with a mixed P/A conformation migrate a little more slowly than the sequences in the antiparallel form. Moreover, by UV-melting we found that the compounds with a mixed P/A conformation, **7214** and **7216**, have a thermal stability, 56.8 and 55.0 °C, respectively, similar to that of **7217** (56.2 °C), which lacks the anthraquinone insertions (Table 1). This suggests that in the oligonucleotide with a mixed P/A conformation, the two anthraquinone insertions, which are placed at opposite ends of the quadruplex, should not promote stabilizing $\pi$-interactions neither with the external G-quartets


DOI: 10.1021/acsmedchemlett.5b00315
ACS Med. Chem. Lett. 2015, 6, 1179−1183

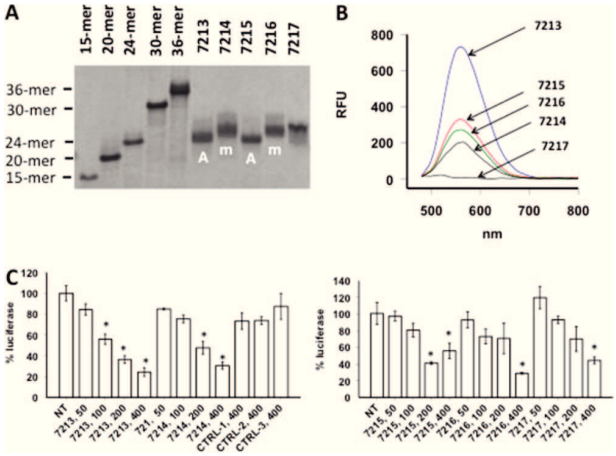



**Figure 3.** (A) Native PAGE of the G4 decoys and of 15-, 20-, 24-, 30-, and 36-mer reference oligonucleotides, *m* = mixed P/A, A = antiparallel. (B) Fluorescence spectra of the decoy oligonucleotides in 50 mM Tris-HCl pH 7.4, 100 mM KCl; Ex = 444 nm, Em = 470–750 nm. (C) Dual luciferase assay with pHRAS-luc, pRL-CMV, and decoy oligonucleotides. The ordinate reports % luciferase, i.e., T/C × 100, where T = firely/*Renilla* in oligonucleotide-treated cells and C = firely/*Renilla* in oligonucleotide-untreated cells. $^{*}P < 0.05$.

nor with the neighboring bases. In contrast, compound **7213** with insertions $H_2$ and $H_4$ in the two lateral loops lying at the same side of the quadruplex (Figure 2A,D), shows a $T_M$ 8 °C higher than that of **7217**. In this oligonucleotide the two anthraquinones are in close proximity and probably stacked on each other. To support this interpretation we performed fluorescence experiments, as the polycyclic anthraquinone chromophore may change its fluorescence quantum yield when it is involved in $\pi$-interactions.[23] When it is excited at 444 nm, it emits at 560 nm (Figure 3B). By fixing the emission of **7214** to 1, all the other compounds showed a higher emission: **7213**, 3.75-fold more intense, **7215**, 1.75-fold, and **7215**, 1.3-fold. Since the $T_M$ of **7214** (56.8 °C), with two anthraquinone at the 5′ and 3′ ends, is similar to the $T_M$ of **7217** (56.2 °C), we jumped to the conclusion that in **7214** the two chromophores are oriented outside the structure and do not promote $\pi$-interactions with the adjacent G-quartets. In contrast, compound **7213** shows a dramatic increase of fluorescence, as the two anthraquinones, inserted in neighboring lateral loops, promote reciprocal $\pi$-interactions. In this case the quadruplex is stabilized by 8 °C compared to **7217**.

In keeping with this interpretation, compound **7215** shows an increase of both $T_M$ (4 °C) and fluorescence (1.75 fold), as the two anthraquinones lying at the same side of the quadruplex can promote $\pi$-interactions, to some extent. Instead, **7216** shows $T_M$ similar to that of **7217**, as the anthraquinones are at opposing ends of the quadruplex (Figure 2D). In a parallel work, we inserted two anthraquinones in the neighboring loops of the thrombin aptamer, which is known to adopt an antiparallel structure.[24] This modified aptamer showed a CD similar to that of **7213** (S3). Molecular modeling studies showed that the two anthraquinones are indeed stacked in a face-to-face manner, as we have hypothesized for the parental **7213** compound (S4).

Next, we interrogated if the designed oligonucleotides behave as decoys and are able to downregulate *HRAS* in human T24 bladder cancer cells. To address this issue, we first carried out a dual luciferase assay with pHRAS-luc (an expressing vector where the whole *HRAS* promoter drives the synthesis of firefly luciferase[10]) and pRL-CMV (a vector in which *Renilla* luciferase is driven by the CMV promoter). T24 cancer cells were treated with increasing amounts (50–400 nM) of G4 oligonucleotides, control oligonucleotides unable to form a G-quadruplex [CTRL-1, 5′-GAGGGAGC-GGCTGC-AGCGCTGGGAG; CTRL-2, 5′-GAAGGAGTG-AGTGAGG-GAGCGGCTGCAG; CRTL-3, 5′-TGCAGCC-GCTCCC-TCACTCACTCCTTCCCT], pHRAS-luc, and pRL-CMV. Figure 3C shows that all the designed G4 oligonucleotides cause the inhibition of firefly luciferase compared to *Renilla* luciferase in a dose–response manner, whereas the three control sequences that did not assume any folded structure, were not effective, even at their highest concentration (400 nM). The result of this reporter assay showed that both types of G4 oligonucleotides, antiparallel and mixed P/A, disrupt the mechanism leading to gene expression, seemingly by sequestering the proteins interacting with *hras*-1. To provide further support for the activity of the designed G4 oligonucleotides, we focused on genomic *HRAS*. We treated T24 cancer cells with 800 nM G4 or control oligonucleotides. Twelve hours after treatment, *HRAS* mRNA in both untreated and oligonucleotide-treated cells was determined by qRT-PCR. The qRT-PCR data showed that all G4 decoys, but not the control oligonucleotides, reduced *HRAS* mRNA compared to those of $\beta$2-microglobulin and HPRT to ∼50% of the control (untreated cells) (S5). We also examined the effect of the G4 decoys on *HRAS* expression by Western blots (Figure 4). In

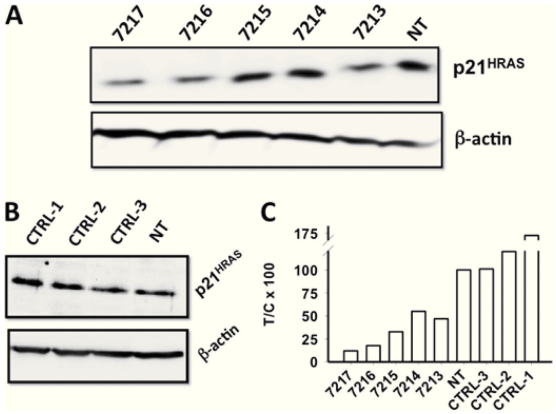



**Figure 4.** Western blot showing that the G4 decoys (A) but not the control oligonucleotides (B) (800 nM) reduce protein $p21^{HRAS}$ in T24 cancer cells. The % inhibition of $p21^{HRAS}$ is shown in the histogram as T/C × 100, where T and C are the $p21^{HRAS}/\beta$-actin ratios in treated and untreated cells, respectively.

accord with the luciferase and qRT-PCR data, all the compounds decreased the level of protein *HRAS*, in particular **7215**, **7216**, and **7217** (to <30% of control), while the control oligonucleotides CTRL 1–3 did not.

Next, we examined the stability of the designed decoy oligonucleotides in a T24 cellular extract, up to 93 h of incubation (Figure 5A). The results showed that wild-type *hras*-1 oligonucleotide undergoes a nearly complete degradation in 93 h. In contrast, all the other compounds appeared more stable, and this provides a key to rationalize their activity. So, there is a correlation between the stability of the G4 oligonucleotides in the cell environment and their capacity to lower luciferase and *HRAS* expression.



DOI: 10.1021/acsmedchemlett.5b00315
*ACS Med. Chem. Lett.* 2015, 6, 1179–1183

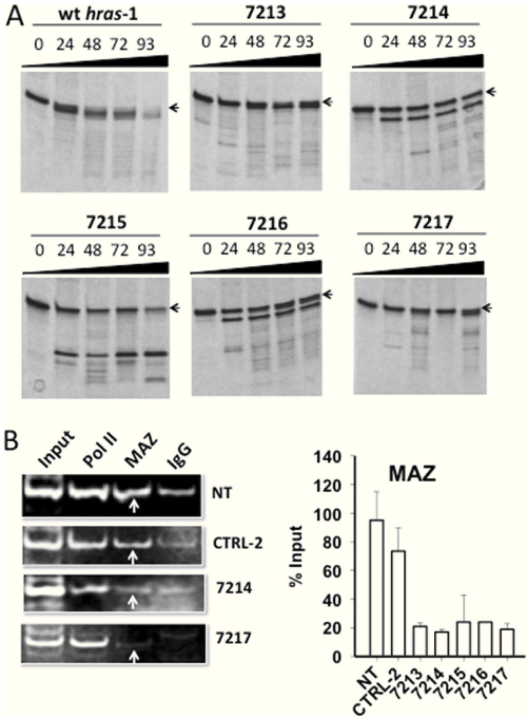



**Figure 5.** (A) Denaturing PAGE showing the stability of the G4 decoys and wild-type *hras*-1 in a T24 cellular extract up to 93 h. (B) ChIP showing that MAZ is associated with the *HRAS* promoter in the untreated (NT) or control oligonucleotide-treated T24 cancer cells (CTRL-2). In contrast, in **7214** and **7217** treated cells, the binding of MAZ to the promoter is attenuated. The product of ChIP-PCR is a 188 bp fragment (S1). The histogram shows levels of MAZ associated with *HRAS* promoter in untreated and control/decoy-treated cells.

As the designed G4 oligonucleotides suppress *HRAS*, we wondered if they really act through a decoy mechanism. To provide evidence for this, we examined by chromatin immunoprecipitation (ChIP) if the G4 decoys attenuate the occupancy of the *HRAS* promoter by proteins essential for transcription. We previously demonstrated that MAZ is a key transcription factor for *HRAS*;[9,10] therefore, we focused our ChIP analysis on this protein, by using an antibody specific for MAZ. A typical ChIP analysis for **7214**, **7217**, and CTRL-2 is reported in Figure 5B. The histogram summarizes the results obtained with all the G4 decoys. The occupancy of the *HRAS* promoter by MAZ in the decoy-treated cells is about 1/3 of the occupancy observed with the untreated or control-treated cells. Note that *hras*-1, being located near to TSS, is significantly occupied by RNA Pol II too (positive control) but not by IgG (negative control).

Furthermore, by an EMSA competition assay we found that all the designed G4 decoys are able to compete with the binding of MAZ to the *hras*-1 quadruplex (S6). So, both ChIP-PCR and EMSA proved that MAZ, under *in vivo* conditions, is associated with the critical G-rich elements of the *HRAS* promoter and that the binding of MAZ to the *HRAS* promoter is attenuated by the G4 decoy oligonucleotides.

As a final point we tested the efficacy of the G4 decoys to inhibit the proliferation of the *HRAS* mutant T24 bladder cancer cells. The cells were transfected with increasing amounts of decoy or control oligonucleotides and incubated for 72 h, before a resazurin assay was carried out. To various extents, all compounds lowered the percentage of viable cells in a dose–response manner. The estimated $IC_{50}$ values (nM) are ~600 for **7213**, ~300 for **7214**, >800 for **7215**, ~800 for **7216**, and >800 for **7217**. In contrast, in noncancer 293 cells the compounds had a weaker effect on cell viability, with $IC_{50} \gg 800$ nM (S7).

We have designed a decoy strategy to inhibit the expression of oncogene *HRAS* in bladder cancer cells. To strengthen the nuclease resistance and the folded conformation, the G4 decoys have been synthesized with LNA modifications and anthraquinone insertions. Luciferase, qRT-PCR, and Western blot assays showed that the decoy oligonucleotide repressed *HRAS* expression, while ChIP and EMSA provided evidence that the molecules actually behave through the postulated decoy mechanism. It might be argued that the decoy strategy lacks specificity because the sequestration of a transcription factor (as MAZ) could have impact on other genes as well. However, bladder cancer cells are addicted to oncogenic *HRAS*, just as pancreatic cancer cells are addicted to mutant *KRAS*.[25] This means that the metabolism of these transformed cells strongly depends on the expression of these oncogenes. Indeed, it has been reported that the *ras* oncogenes remodulate the metabolic pathways of cancer cells to fuel their higher proliferation rate.[25] So, cancer cells are more responsive to ras gene expression than normal cells. In other words, the inhibition of *HRAS* in *HRAS* mutant bladder cancer cells is less tolerated than in noncancer 293 cells bearing wild-type *HRAS*. The concept of oncogene addiction for targeted anticancer therapy has been recently illustrated.[26]

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsmedchemlett.5b00315.

Materials and Methods, oligonucleotides synthesis, CD spectra, qRT-PCR, EMSA, and cell viability assays (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

*E-mail: luigi.xodo@uniud.it. Fax: +39.043249430.

### Present Address

[‖]Department of Chemistry, Faculty of Science, Benha University, Benha 13518, Egypt.


### Author Contributions

[§]G.M. and A.S.G. have contributed equally to this work. The manuscript was written by L.X. All authors have given approval to the final version of the manuscript.


### Funding

AIRC (Italian association for cancer research), IG2013, project Code 14301


### Notes

The authors declare no competing financial interest.

## ■ ABBREVIATIONS

ChIP, chromatin immune precipitation; HRAS, Harvey ras gene; MAZ, myc associated zinc-finger protein; LNA, locked nucleic acid

## ■ REFERENCES


(1) Ding, L.; Getz, G.; Wheeler, D. A.; et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **2008**, *455*, 1069–1075.




DOI: 10.1021/acsmedchemlett.5b00315
*ACS Med. Chem. Lett.* 2015, 6, 1179–1183

(2) Porter, A. C.; Vaillancourt, R. R. Tyrosine kinase receptor-activated signal transduction pathways which lead to oncogenesis. *Oncogene* **1998**, *17*, 1343−1352.
(3) Schubbert, S.; Shannon, K.; Bollag, G. Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer* **2007**, *7*, 295−308.
(4) Vageli, D.; Kiaris, H.; Delakas, D.; Anezinis, P.; Cranidis, A.; Spandidos, D. A. Transcriptional activation of H-ras, K-ras and N-ras proto-oncogenes in human bladder tumors. *Cancer Lett.* **1996**, *107*, 241−247.
(5) Theodorescu, D.; Cornil, I.; Fernandez, B. J.; Kerbel, R. Overexpression of normal and mutated forms of HRAS induces orthotopic bladder invasion in a human transitional cell carcinoma. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 9047−9051.
(6) Mo, L.; Zheng, X.; Huang, H. Y.; Shapiro, E.; Lepor, H.; Cordon-Cardo, C.; Sun, T. T.; Wu, X. R. Hyperactivation of Ha-ras oncogene, but not Ink4a/Arf deficiency, triggers bladder tumorigenesis. *J. Clin. Invest.* **2007**, *117*, 314−325.
(7) Yafi, F. A.; North, S.; Kassouf, W. First- and second-line therapy for metastatic urothelial carcinoma of the bladder. *Curr. Oncol.* **2011**, *18*, e25−e34.
(8) Baines, A. T.; Depeng, X.; Der, C. J. Inhibition of Ras for cancer treatment: the search continue. *Future Med. Chem.* **2011**, *3*, 1787−1808.
(9) Membrino, A.; Cogoi, S.; Pedersen, E. B.; Xodo, L. E. G4-DNA formation in the HRAS promoter and rational design of decoy oligonucleotides for cancer therapy. *PLoS One* **2011**, *6*, e24421.
(10) Cogoi, S.; Shchekotikhin, A. E.; Xodo, L. E. HRAS is silenced by two neighbouring G-quadruplexes and activated by MAZ, a zinc-finger transcription factor with DNA unfolding property. *Nucleic Acids Res.* **2014**, *42*, 8379−8388.
(11) Cogoi, S.; Zorzet, S.; Rapozzi, V.; Géci, I.; Pedersen, E. B.; Xodo, L. E. MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic Acids Res.* **2013**, *41*, 4049−4064.
(12) Sun, X.; Zhang, J. STAT3 Decoy ODN Therapy for Cancer. *Methods Mol. Biol.* **2015**, *1317*, 167−83.
(13) Kim, K. H.; Lee, W. R.; Kang, Y. N.; Chang, Y. C.; Park, K. K. Inhibitory effect of nuclear factor-κB decoy oligo-deoxynucleotide on liver fibrosis through regulation of the epithelial-mesenchymal transition. *Hum. Gene Ther.* **2014**, *25*, 721−729.
(14) Chan, K. S.; Sano, S.; Kiguchi, K.; Anders, J.; Komazawa, N.; Takeda, J.; DiGiovanni, J. Disruption of Stat3 reveals a critical role in both the initiation and the promotion stages of epithelial carcinogenesis. *J. Clin. Invest.* **2004**, *114*, 720−728.
(15) Sen, M.; Paul, K.; Freilino, M. L.; Li, H.; Li, C.; Johnson, D. E.; Wangm, L.; Eiseman, J.; Grandis, J. R. Systemic administration of a cyclic signal transducer and activator of transcription 3 (STAT3) decoy oligonucleotide inhibits tumor growth without inducing toxicological effects. *Mol. Med.* **2014**, *20*, 46−56.
(16) Kim, K. H.; Park, J. H.; Lee, W. R.; Park, J. S.; Kim, H. C.; Park, K. K. The inhibitory effect of chimeric decoy oligodeoxynucleotide against NF-κB and Sp1 in renal interstitial fibrosis. *J. Mol. Med. (Heidelberg, Ger.)* **2013**, *91*, 573−586.
(17) De Rosa, G.; De Stefano, D.; Laguardia, V.; Arpicco, S.; Simeon, V.; Carnuccio, R.; Fattal, E. Novel cationic liposome formulation for the delivery of an oligonucleotide decoy to NF-kappaB into activated macrophages. *Eur. J. Pharm. Biopharm.* **2008**, *70*, 7−18.
(18) Cogoi, S.; Zorzet, S.; Rapozzi, V.; Géci, I.; Pedersen, E. B.; Xodo, L. E. MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic Acids Res.* **2013**, *41*, 4049−4064.
(19) Campbell, M. A.; Wengel, J. Locked vs. unlocked nucleic acids (LNA vs. UNA): contrasting structures work towards common therapeutic goals. *Chem. Soc. Rev.* **2011**, *40*, 5680−5689.
(20) Ndlebe, T.; Schuster, G. B. Long-distance radical cation transport in DNA: horizontal charge hopping in a dimeric quadruplex. *Org. Biomol. Chem.* **2006**, *4*, 4015−4021.
(21) Gouda, A. S.; Amine, M. S.; Pedersen, E. B. Synthesis of new DNA G-quadruplex constructs with anthraquinone insertions and their anticoagulant activity. *Helv. Chim. Acta*, in press.
(22) Randazzo, A.; Spada, G. P.; da Silva, M. W. Circular dichroism of quadruplex structures. *Top. Curr. Chem.* **2012**, *330*, 67−86.
(23) Suzuki, I.; Nakayama, C.; Ui, M.; Hirose, K.; Yamauchi, A. Greater fluorescence from styrylpyridonium dye upon complexation with cyclodextrin derivatives through π-π interactions. *Anal. Sci.* **2007**, *3*, 249−251.
(24) Mao, X. A.; Gmeiner, W. H. NMR study of the folding-unfolding mechanism for the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). *Biophys. Chem.* **2005**, *113*, 155−160.
(25) Ying, H.; Kimmelman, A. C.; Lyssiotis, C. A.; et al. Oncogenic Kras maintains pancreatic tumours through regulation of anabolic glucose metabolism. *Cell* **2012**, *149*, 656−670.
(26) Torti, D.; Truscolino, L. Oncogene addiction as a foundational rationale for targeted anti-cancer therapy: promises and perils. *EMBO Mol. Med.* **2011**, *3*, 623−636.


DOI: 10.1021/acsmedchemlett.5b00315
*ACS Med. Chem. Lett.* 2015, 6, 1179−1183

# SUPPORTING INFORMATION

## NUCLEIC ACID TARGETED THERAPY: G4 OLIGONUCLEOTIDES DOWNREGULATE *HRAS* IN BLADDER CANCER CELLS THROUGH A DECOY MECHANISM

Giulia Miglietta,[†] Alaa S. Gouda,[&] Susanna Cogoi,[†] Erik B. Pedersen[&] and Luigi E. Xodo[†]

[†] Department of Medical and Biological Sciences, University of Udine, 33100 Italy; [&] Nucleic Acid Center, Institute of Physics and Chemistry, University of Southern Denmark, DK-5230 Odense M, Denmark.

### S1: MATERIALS AND METHODS

#### Oligonucleotides

Anthraquinone-modified oligonucleotides have been synthesized as previously described (Alaa S. Gouda, A.S., Amine, M.S., Pedersen, E.B. Synthesis of New DNA G-Quadruplex Constructs with Anthraquinone Insertions and Their Anticoagulant Activity. *Helv. Chim. Acta.,* in press). The synthesized oligonucleotides have been purified by HPLC and the correct molecular weight of the oligonucleotides was verified by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry analysis (Table 1). Equivalence between calculated and measured molecular weights were found (difference < 0.05 %). The wild type DNA oligonucleotide, DNA phosphoramidite monomers, solid supports and additional reagents were purchased from Sigma-Aldrich or Glen Research**.**

#### Synthesis, purification and molecular weight of decoy oligonucleotides.

The solid phase synthesis of the decoy oligonucleotides was carried out on a PerSeptive Biosystems expedite 8909 automated DNA/RNA synthesizer in 1.0 μmol scale on 500Å

CPG supports. A standard cycle protocol was applied for 1,4-substituted anthraquinone phosphoramidite using 5-[3,5-bis(trifluoromethyl)phenyl]-*1H*-tetrazole (0.25 M, in dry acetonitrile) as an activator, followed by insertion via hand-coupling into the growing oligonucleotides chain during coupling time (15 min). Stepwise coupling efficiencies, determined by a spectrophotometric DMT assay, were >99% for the synthesized phosphoramidite. Removal of Nucleobase protecting groups and cleavage from solid support were carried out under standard conditions (1ml of 32% aqueous ammonia, 12 h, 55 °C). The resulting oligonucleotides were purified by DMTr-ON reversed phase HPLC (RP-HPLC) using the Waters system 600 equipped with column (XBridge OST C18, 19x1000mm, 5µm + precolumn: XBridge 10x10mm, 5µm, temperature column oven: 50 °C). Elution was performed starting with an isocratic hold of buffer *A* for 2min followed by a linear gradient to 70 % buffer *B* over 17 min at a flow rate of 5ml/min (Buffer *A*: 0.05M triethylammonium acetate in Milli-Q water, pH 7.4; Buffer *B*: 75 % $CH_3CN$ / 25 % Buffer *A*). All oligonucleotides were detrityled, after removing solvents under nitrogen flow, by treatment with 100 µL of 80% aqueous solution of acetic acid for 30 min, 100 µL double filtered water, then quenched or desalted with an aqueous solution of sodium acetate (3M, 15 µL) and aqueous sodium perchlorate (5 M, 15 µL) followed by cold acetone (1ml). The resulting mixture was precipitated at -20 °C for 1 h. After centrifugation (13000 rpm,10 min, 4 °C), the supernatant was removed and the pellet washed with cold acetone (2x 1ml),dried for 30 min under nitrogen flow, and dissolved in MilliQ water (1000 µl). The composition of the oligonucleotides was verified by MALDI-TOF (Matrix assisted laser desorption/ionization-time of flight) analysis on Microflex *LT* from Bruker Daltonics (MALDI-LIFT system MS instrument in $ES^+$ mode with HPA-matrix (10 mg 3-hydroxypicolinic acid in 50 mM ammoniumcitrate / 70 % acetonitrile) matrix. The purity of the final quadruplexes was more than 90%, recorded by analytical (IE-HPLC) on a Merck Hitachi La-Chrom system (DNAPac PA100, 13µm, 4mm x 250 mm, analytical column) heated to 60 °C. Elution was performed with an isocratic hold of buffer B (10 %), starting from 2 min hold on 2 % Buffer A in Milli-Q water (solvent A), followed by a linear gradient to 30 % buffer A in 23 min at a flow rate of 1.1 mL/min (buffer A: 1.0 M sodium perchlorate; buffer B: 0.25 M Tris-Cl, pH 8.0; solvent A: Milli-Q water).

Mass Spectrometry spectra**.** Electrospray ionization high-resolution mass spectra (ESI-HRMS) were performed on PE SCIEX API Q-Star Pulsar Mass Spectrometer. For accurate

ion mass determinations, the ($M+H^+$) or ($M+Na^+$) ion was peak matched by calibration with NaI.

**Cell culture and proliferation assay**

T24 urinary bladder cancer and human embryonic kidney 293 cells were maintained in exponential growth in Dulbecco's Modified Eagle's Medium (DMEM) containing 100 U/mL penicillin, 100 mg/mL streptomycin, 20 mM L-glutamine and 10% foetal bovine serum (Euroclone, Milan, Italy). For metabolic assays, the cells were seeded [5000 (T24) or 20000 (293) cells/well] the day before G4-decoy treatment in a 96-well plate. Metabolic activity of untreated or oligonucleotide-treated (50-800 nM) cells was measured by a resazurin assay following standard procedures.

**Dual luciferase assays**

Transfection was performed by mixing vector (250 ng/well) pHRAS-luc with 10 ng/well of control plasmid pRL-CMV where *Renilla* luciferase was driven by the CMV promoter using jet-PEI (Polyplus) as a transfecting reagent. Vector pHRAS-luc expressed firefly luciferase under the control of wild-type *HRAS* promoter. Firefly luciferase in cell lysates was measured and normalized for *Renilla* luciferase. Luciferase assays were performed 48 after transfection with Dual-Glo Luciferase Assay System (Promega) following the supplier instructions.

Samples were read on a Turner Luminometer and the relative luminescence expressed as (T/C x 100) where T=*Firefly* luciferase/*Renilla* luciferase in treated cells and C=*Firefly* luciferase/*renilla* luciferase in untreated cells. Each transfection was performed in triplicate.

Plasmid pHRAS-luc was previously constructed in our laboratory (Ref. 9, text).

**RNA extraction and qRT-PCR determination of *HRAS* mRNA**

T24 bladder cells have been plated in 96-well plate (8000 cells/well). 12 h after treatment, RNA was extracted by using iScript $^{TM}$ RT-qPCR sample preparation reagent (BioRad, USA).

For cDNA synthesis, 1.25 µl of RNA was heated at 70 °C and placed in ice. The solution was added with 7.5 µl of a mix containing (final concentrations) 1×buffer; 0.01 M DTT; 1.6 µM

primer dT [MWG Biotech, Ebersberg, Germany; d(T)16]; and 1.6 µM random primers; 0.4 mM dNTPs solution containing equimolar amounts of dATP, dCTP, dGTP and dTTP (Euroclone, Pavia, Italy); 0.8 U/µl RNAse OUT; 8 U/µl of M-MLV reverse transcriptase (Life Technologies, Monza, Italy). The reactions were incubated for 1 h at 37°C and stopped by heating at 95°C for 5 min. As a negative control the reverse transcription reaction was performed with a sample containing DEPC water.

Real-time PCR multiplex reactions were performed with 1xKapa Probe fast qPCR kit (KAPA Biosystems, Wilmington, MA, USA) for *HRAS* and housekeeping genes β2-microglobulin and HPRT, 2.2 µl of cDNA and primers/probes at the following concentrations: for *HRAS* accession n. NM 001130442: probe FAM-CAATGACCACCTGCTTCCG-BHQ1 (from 309 to 327, 260 nM), primer sense 5'-GCTGATCCAGAACCATTT (from 254 to 271, 340 nM), primer antisense 5'-GTATCCAGGATGTCCAAC (from 344 to 361, 260 nM). For β2-microglobulin accession n. NM_004048 probe ROX-TATGCCTGCCGTGTGAACC-BHQ2 (from 352 to 370, 60 nM), primer sense 5'-CCCCACTGAAAAAGATGA (from 333 to 350, 100 nM), primer antisense 5'-CCATGATGCTGCTTACAT (from 415 to 432, 100 nM). For HPRT accession n. NM_000194 probe 5'-Cy5-CTTGCGACCTTGACCATCTT-BHQ2 (from 633 to 652, 180 nM), primer sense 5'-CTTGATTGTGGAAGATATAATTG (from 557 to 575, 210 nM), primer antisense 5'-TATATCCAACACTTCGTGG (from 672 to 690, 230 nM). The PCR cycle was: 3 min at 95 °C, 50 cycles 10 s at 95 °C, 60 s at 58 °C. PCR reactions were carried out with a CFX-96 real-time PCR apparatus controlled by an Optical System software (version 3.1) (BioRad Laboratories, CA, USA). *HRAS* mRNA was normalized with the two housekeeping genes.

**Western blots**

Total protein lysates in Laemmli buffer (40 µg), from T24 cells untreated or oligonucleotide-treated, were electrophoresed on 12 % SDS-PAGE and transferred to a nitrocellulose membrane at 70 V for 2 h. The filter was blocked for 1 h with 5% nonfat dry milk solution in PBS 0.05 % Tween (Sigma-Aldrich, Milan, Italy) at room temperature. The primary antibodies used are: anti-actin (IgM mouse 1:10.000 ab-1 Calbiochem), anti-Hras (IgG rabbit 1:500 ab32417 Abcam). Membranes with the samples were overnight incubated at 4 °C with primary antibodies. The filters were washed with a 0.05 % Tween in PBS and subsequently incubated for 1 h with the secondary antibodies horseradish peroxidase conjugated: anti-rabbit IgG, anti-mouse IgM, diluted 1:5000 (Calbiochem). For the detection of the proteins

we used Super Signal®West PICO, and FEMTO (Thermo Fisher Scientific Pierce). Exposure length depends on the antibodies used.

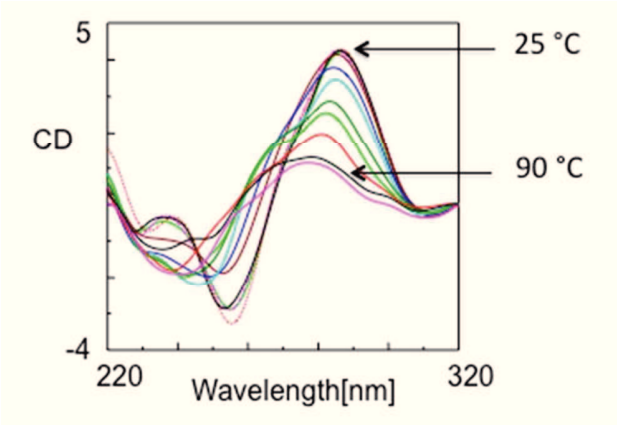
**PAGE and Stability of oligonucleotides in cellular environment**

Native PAGE was performed in a 20 % polyacrylamide gel in TBE. The oligonucleotides (10 μM) were stained with "stains all" dye. For the stability in the cellular extract, the designed G4oligonucleotides and wild type *hras*-1 (3 μM) have been incubated up to 93 h in a cellular extract obtained from T24 cancer cells as previously described (Nucleic Acids Res. 2008, 36, 3765-80). After incubation the oligonucleotides have been run in a denaturing 20 % polyacrylamide gel (7 M urea, 1 x TBE).

**Chromatin immunoprecipitation**

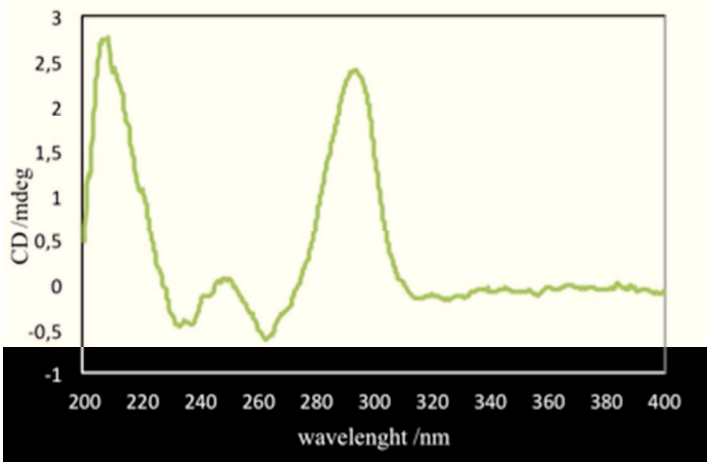
T24 urinary bladder cancer cells ($1.0 \times 10^6$) were cultured overnight in 6-cm diameter plates up to about 80% confluency and fixed in 1% formaldehyde in PBS for 5 minutes at room temperature to crosslink proteins to DNA. Chromatin immunoprecipitation assays were performed using the ChIP-IT$^{TM}$ Express kit (Active Motif, Rixensart, Belgium). The cells were lysed in 100 mM NaCl, 20 mM Tris-HCl, pH 7.4, 1 mM EDTA, 0.5 % NP-40, 0.5 % Na-deoxycholate, 0.1 SDS and sonicated to fragment chromosomal DNA into ~ 500 bp (8 min, 20 kHz, Bioruptor, Diagenode, NY). Sheared chromatin (6 μg in each sample) was diluted and incubated overnight, at 4 °C, with 0.5 μg of each specific antibody (Ab) and pulled down with Protein G magnetic beads. The Abs used are MAZ, (Santa Cruz Inc., USA), negative control mouse IgG (Active Motif), positive control RNA Pol II mouse monoclonal Ab (Active Motif). After incubation, the mixtures were spinned and the beads washed 3-times with kit washing buffers. The chromatin was treated with kit elution buffer for 15 min at RT on a shaker. We then used proteinase K (37 °C) to reverse the crosslinking and after 1 h we stopped the reaction with a proteinase K stop solution. The DNA recovered was amplified by PCR, using the following primers (accession number J00277): (i) 5'-GGCTCCTGACAGACGGG (304-320; *hras*-1for) and 5'-GCATGGGCTCCGTCC (477-491; *hras*-1rev) giving a 188 bp product. The amplified product were separated by 8 % acrylamide gel in TBE and quantified with a Gel–DOC apparatus (Bio-Rad Laboratories, CA, USA). The data have been normalized by input.

**S2: CD spectra of sequence *hras*-1 in 100 mM Tris-HCl, pH 7.4, 100 mM KCl, as a function of temperature.** The strong positive ellipticity at 290 nm and the negative ellipticity at near 260 nm are typical of an antiparallel quadruplex.

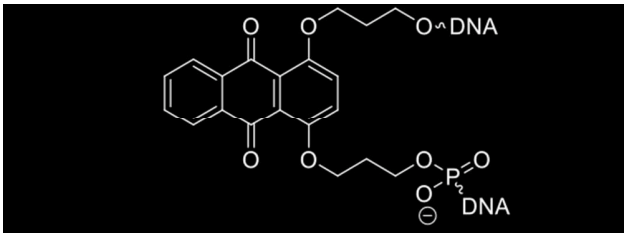



**S3: (A) CD spectrum of a modified thrombin aptamer bearing two anthraquinone insertions (B) in the lateral neighbouring loops as indicated in panel (C).**

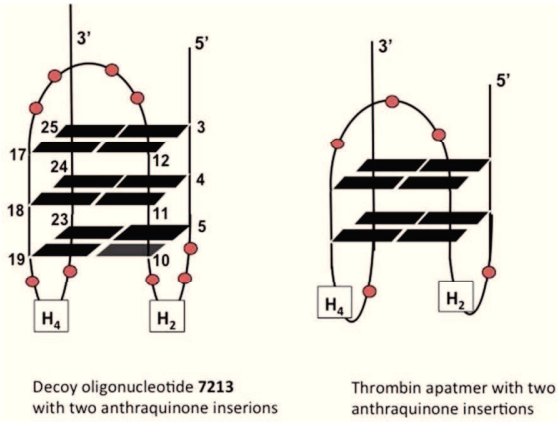
(A)



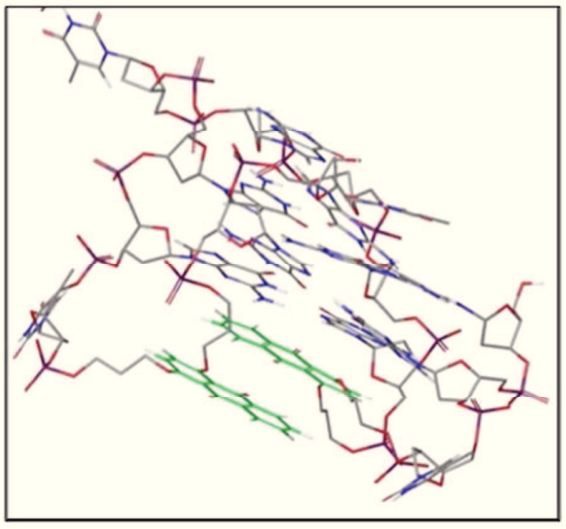

(B)



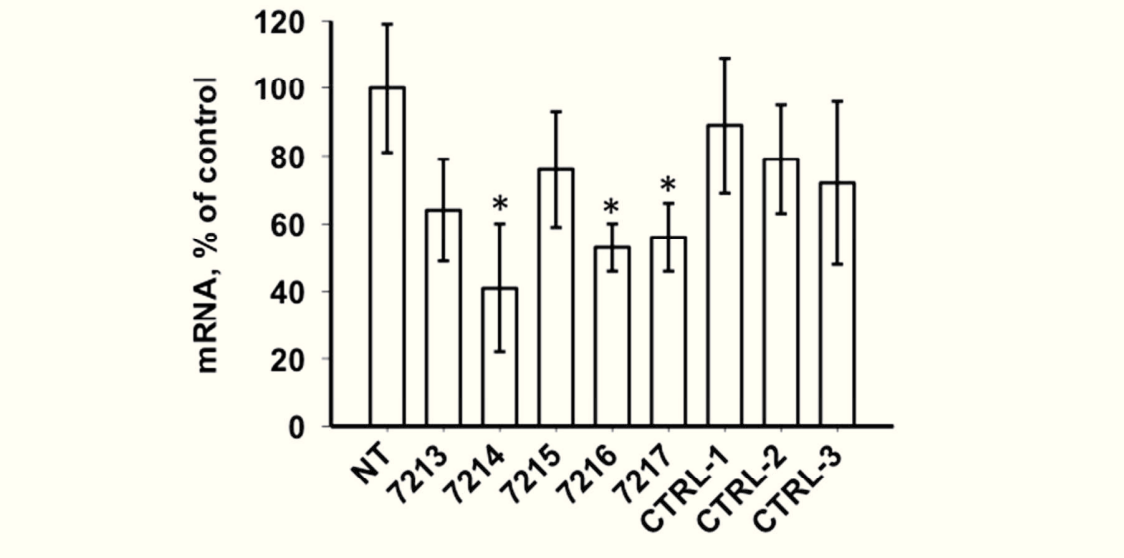

(C)



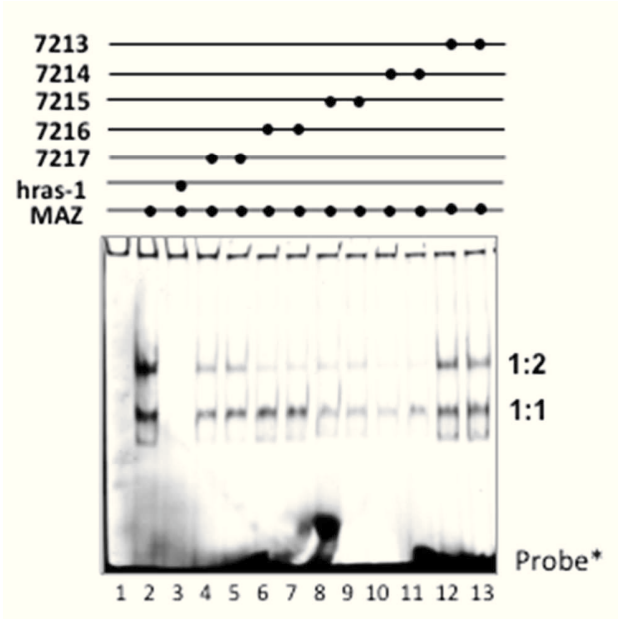

**S4: Molecular model of thrombin aptamer with two anthraquinone insertions in the two lateral loops.** A modified AMBER* force field in Macro Model with (Maestro v9.2) molecular modeling was utilized to generate representative low-energy structures of the anthraquinone-modified 15mer thrombin aptamer. Note that the two anthraquinone chromophores are arranged in a face-to-face stacking manner that increases the stability of the quadruplex. A similar structure is expected for the parental **7213** anthraquinone quadruplex.



**S5: Determination of the level of HRAS mRNA by quantitative real-time PCR.** The assay has been conducted after the G4-decoys (800 nM) have been incubated for 12 h with T24 cancer cells. Asterisk indicates P< 0.05.



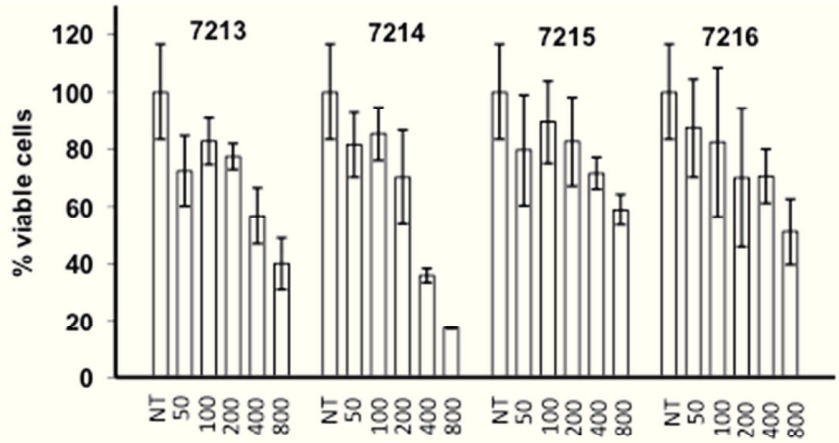

**S6: EMSA competition assay.** The binding between MAZ and the G4-decoy oligonucleotides has been investigated by EMSA competition assay. Quadruplex *hras*-1 labeled with dy781 (30 nM), was incubated with 50- or 100-fold excess of G4-decoy oligonucleotide and 3 μg of recombinant MAZ-GST for 30 min at 25 °C, in 50 mM Tris-acetate, pH 7.4, 50 mM KCl, 1 mM EDTA, 2.5ng/μl Salmon sperm. After incubation, the reaction mixtures were loaded in 5% PAGE 1xTBE, thermostated at 20°C. After running the gel was analyzed by Odyssey CL x scanner /ImageStudio Software (Li-Cor Biosciences). Lane 2 shows the binding of MAZ to hras-1 forming, as previously reported, a 1:1 and 1:2 complexes. When the competitor is hras-1 itself (100-fold), the binding is completely abrogated, as expected (lane 3). Instead, when we used as competitors the G4 decoys, we observed a different extent of competition, although all the oligonucleotides showed the capacity to compete the binding of MAZ to the promoter hras-1 sequence. Note that in the absence of competitor the 1:2 complex is more abundant than complex 1:1. In contrast, in the presence of the competitor, the 1:1 complex is instead more abundant, ss the G4 decoys are found to attenuate the binding of MAZ to the promoter.

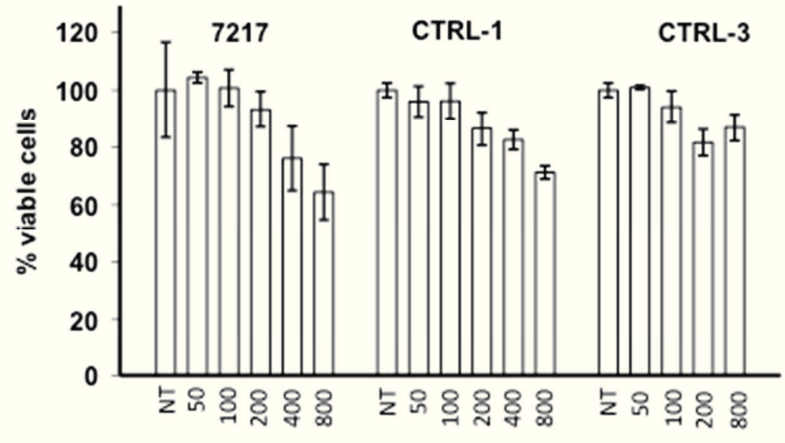


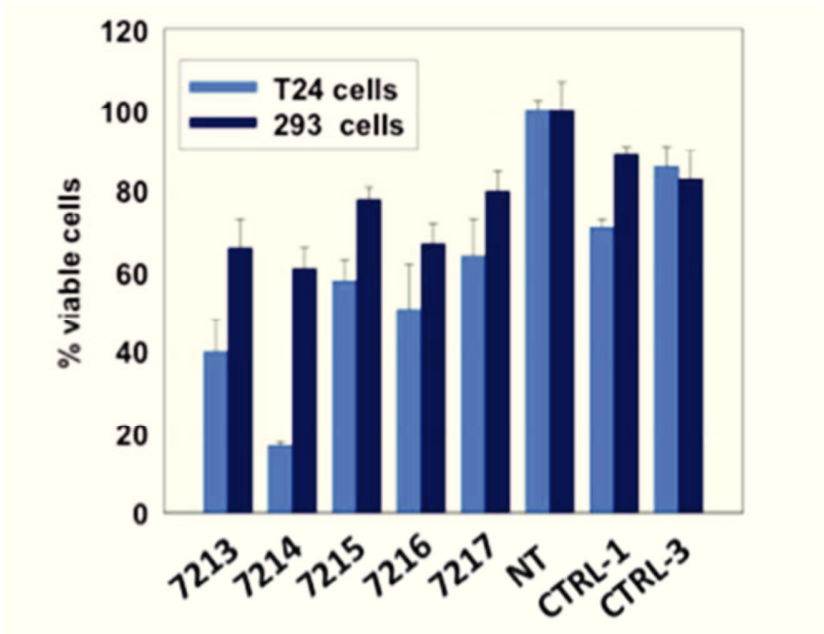

Lane 1: hras-1 labelled to dy781 (probe*, 30 nM) ; lane 2: probe and MAZ (3 μg); lane 3: as lane 2 plus 100-fold unlabelled hras-1; lanes 4-5, 6-7, 8-9, 10-11, 12-13 as in lanes 2 plus 50- and 100-fold G4 decoy 7217, 7216, 7215, 7214 and 7213, respectively.

**S7**: **Viability assays.** (Panel 1 and 2 from top) Percentage of viable cells determined by a resazurin assay. Human T24 bladder cancer cells have been treated with increasing amounts (0-50-100-200-400-800 nM) of decoy or control oligonucleotides for 72 h. After incubation a resazurin assay was carried out. The % viable cells is given by T/C x 100, where T is the fluorescence of oligonucleotide-treated cells and C is the fluorescence of untreated cells. The data are the average of three experiments, the standard deviations are also reported;

(Panel 3) Percentage of viable T24 and 293 cells after incubation for 72 h with 800 nM decoy and control oligonucleotides. T24 cancer cells are *HRAS*-mutated, while 293 noncancer cells bear wild-type *HRAS*. The percentage of viable 293 cells varies from 61 % to 80%, while the % of viable T24 cells varies from 17 % to 64 %.


7213
7214
7215
7216
% viable cells
120
100
80
60
40
20
0
NT
50
100
200
400
800


7217
CTRL-1
CTRL-3
% viable cells
120
100
80
60
40
20
0
NT
50
100
200
400
800

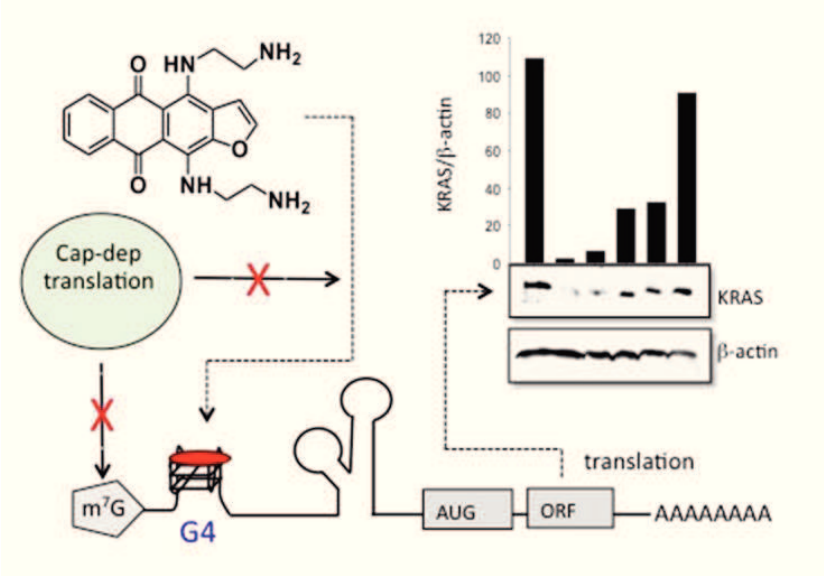

T24 cells
293 cells
% viable cells
120
100
80
60
40
20
0
7213
7214
7215
7216
7217
NT
CTRL-1
CTRL-3

**Section C:**

# RNA G-quadruplexes in Kirsten ras (KRAS) oncogene as targets for small molecules inhibiting translation

G. Miglietta, S. Cogoi, J. Marinello, G. Capranico, AS. Tikhomirov, A. Shchekotikhin and LE. Xodo

*Journal of Medicinal Chemistry* 60 (23), 9448–9461 (2017)




Graphic contents of the article

Journal of Medicinal Chemistry

Article

Cite This: J. Med. Chem. 2017, 60, 9448−9461

pubs.acs.org/jmc

# RNA G-Quadruplexes in Kirsten Ras (*KRAS*) Oncogene as Targets for Small Molecules Inhibiting Translation

Giulia Miglietta,[†] Susanna Cogoi,[†] Jessica Marinello,[‡] Giovanni Capranico,[‡] Alexander S. Tikhomirov,[§] Andrey Shchekotikhin,[§] and Luigi E. Xodo*[,†]

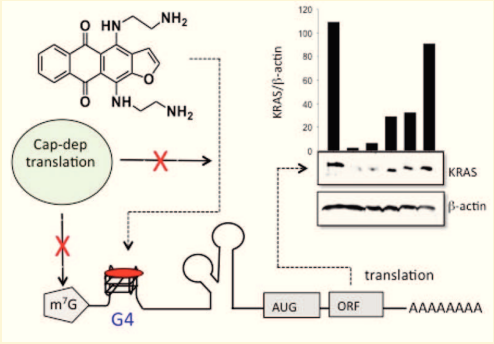
[†]Department of Medicine, Biochemistry Laboratory, University of Udine, 33100 Udine, Italy

[‡]Department of Pharmacy and Biotechnology, University of Bologna, 40100 Bologna, Italy

[§]Gause Institute of New Antibiotics, 119021 Moscow, Russia

Supporting Information

**ABSTRACT:** The human *KRAS* transcript contains a G-rich 5′-UTR sequence (77% GC) harboring several G4 motifs capable to form stable RNA G-quadruplex (RG4) structures that can serve as targets for small molecules. A biotin−streptavidin pull-down assay showed that 4,11-bis(2-aminoethylamino)anthra[2,3-*b*]furan-5,10-dione (**2a**) binds to RG4s in the *KRAS* transcript under low-abundance cellular conditions. Dual-luciferase assays demonstrated that **2a** and its analogue 4,11-bis(2-aminoethylamino)anthra[2,3-*b*]thiophene-5,10-dione (**2b**) repress translation in a dose-dependent manner. The effect of the G4-ligands on Panc-1 cancer cells has also been examined. Both **2a** and **2b** efficiently penetrate the cells, suppressing protein p21KRAS to <10% of the control. The *KRAS* down-regulation induces apoptosis together with a dramatic reduction of cell growth and colony formation. In summary, we report a strategy to suppress the *KRAS* oncogene in pancreatic cancer cells by means of small molecules binding to RG4s in the 5′-UTR of mRNA.




## INTRODUCTION

The three ras genes (*HRAS*, *KRAS*, and *NRAS*) encode for highly homologous (83−90% sequence identity) GTPases of 21 kDa that cycle between an active GTP-bound and an inactive GDP-bound state.[1,2] This cycling is mediated by guanine nucleotides exchange factors (GEFs) and by GTPase activating proteins (GAPs).[1,3] In the GTP-bound state, the p21RAS protein interacts with downstream effectors, activating specific cellular processes including proliferation, survival, and differentiation.[3,4] Mutations in the ras genes are estimated to be present in ∼30% of all human cancers.[5] However, in pancreatic ductal adenocarcinoma (PDAC), *KRAS* is mutated in ∼95% of patients.[6,7] The mutant alleles carry a single missense point mutation in exon 1, codon 12, 13, or 61, which impairs GAP-mediated GTP-to-GDP hydrolysis. This results in an aberrant protein that is locked into the activated state, transmitting constitutively signals for proliferation to the nucleus.[8] According to recent studies, mutations in the *KRAS* gene can be seen as primary genetic lesions that initiate the malignant transformation of pancreatic cells.[9,10] Progression to invasive PDAC occurs through a stepwise accumulation of other genetic lesions, in particular those causing the inactivation of tumor suppressor genes.[11]

Recent studies have demonstrated that *KRAS* is essential for the maintenance of PDAC as it reprograms the metabolism of glucose and glutamine to fuel a high proliferation rate.[9,12,13] The dependence of metabolic pathways on specific oncogenes has led to the concept of "oncogene addiction", which means that, although cancer cells may depend on a number of genetic aberrations, they often develop a dependency on a particular oncogene.[14,15] Considering the central role played by *KRAS* on the pathogenesis of PDAC, *KRAS* is considered a crucial target for anticancer drugs. However, despite more than two decades of research, up to now no antiras drugs have reached the clinic, creating the impression that ras genes may be "undruggable".[16,17] Recently, the design of new inhibitors binding directly to protein p21RAS has fueled research in this direction.[18] Other strategies that are being pursued use drugs that inhibit either the association of p21RAS to the membrane or the activity of downstream pathways.[19−22] In our laboratory, we developed an alternative anti *KRAS* strategy by focusing on two targets for small molecules: (i) a G4-motif located upstream of the transcription start site (TSS) (−320/−306), which is recognized by essential transcription factors (MAZ, Ku70, PARP-1, and hnRNP A1);[23] (ii) G4-motifs located in the 5′-untranslated region (5′-UTR) of the *KRAS* transcript.[24,25] Previous studies have shown that the presence of G4 motifs in 5′-UTR of mRNA inhibits translation, on the basis of luciferase assays.[26,27] In this study, we demonstrate by a streptavidin−biotin pull-down assay that small molecules bind to RNA G-quadruplexes (RG4s) formed in the 5′-UTR of low-abundant

**Received:** May 1, 2017
**Published:** November 15, 2017

ACS Publications © 2017 American Chemical Society



DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

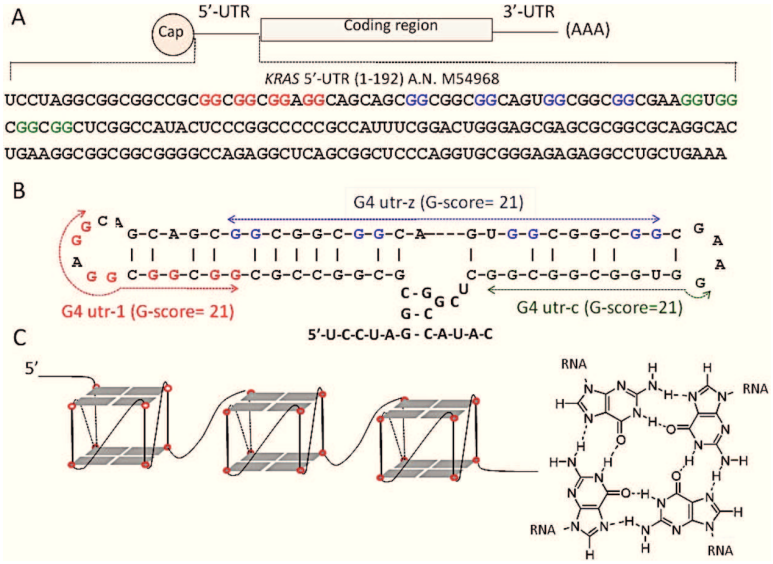



**Figure 1.** (A) Primary structure of the 5′-UTR of the human *KRAS* transcript. The sequence is very rich in guanines (77% GC) and contains 33 GG-runs. (B) Predicted secondary structure of the first 80 nt of 5′-UTR (s-80) given by Mfold.[48] In s-80 there are three putative nonoverlapping G4 motifs *utr*-1, *utr*-z, and *utr*-c, indicated with colored arrows on the stem-loop structure. According to QGRS, the G4 motifs have a G-score = 21. (C) Cartoon showing the three nonoverlapping G4 RNAs and the structure of a G-tetrad.

**Table 1. G4 Motifs in s-80, the First 80 nt in the *KRAS* 5′-UTR mRNA**

| | | position sequence (5′→3′) | G/C[a] | G-score | $T_M$[b] | G4 RNA $\Delta G^c$ (kcal/mol) |
|---|---|---|---|---|---|---|
| *utr*-1 | 16−30 | GCGGCGGCGGAGGCA | 4 | 21 | 53 | −4.9 |
| *utr*-z | 36−55 | GGCGGCGGCAGUGGCGGCGG | 2.6 | 21 | 64 | −3.7 |
| *utr*-c | 59−71 | AGGUGGCGGCGGC | 4 | 21 | 52 | −5.6 |

[a]G/C ratio of the G4 motifs (>1.5 favors G4 over loop-hairpin).[37] [b]$T_M$ obtained in 50 mM Na-cacodylate, pH 7.4, 100 mM KCl. [c]$\Delta G$s (298 K) obtained from UV-melting curves, using a two-state model.

cellular *KRAS* transcripts. These small molecules, anthrafurandiones (ATFD) and anthrathiophenediones (ATPD), suppress luciferase from specific vectors as well as p21RAS in pancreatic Panc-1 cancer cells. ATFD and ATPD strongly induce apoptosis and reduce the metabolic activity and colony formation of Panc-1 cells carrying mutant *KRAS* G12D. Our results demonstrate that ATPD and ATFD are promising therapeutic drugs to suppress oncogenic *KRAS* in pancreatic cancer cells, through their binding to RG4s located in the 5′-UTR of mRNA.

## RESULTS AND DISCUSSION

Since Hurley and co-workers published their seminal study on the transcription of *CMYC* controlled by a regulatory mechanism involving a G4 DNA,[28] a great number of papers on this unusual structure has been reported in the literature.[29−35] The biological function of G4 DNA is rather complex: some studies indicate that it behaves as a transcriptional repressor,[28,32,35] others that G4 DNA is associated with transcriptionally active chromatin.[36] Whatever its function is, G4 DNA is an attractive target for therapeutic small molecules.[32] However, bioinformatics analyses showed that the transcript of certain oncogenes contains a 5′-UTR rich in guanines that can form G4 RNA structures (RG4s).[26,27,37−40] The 5′-UTR region of mRNA plays an essential role in the initiation of translation. This occurs through three steps: (i) association of the 7-methylguanosine cap with the 43 S initiation complex; (ii) scanning of complex 43 S along 5′-UTR up to codon AUG; (iii) assembly of a larger 80 S ribosome that proceeds with protein synthesis. The process is normally regulated by eukaryotic translation initiation factors (eIFs). In addition, certain transcripts use a cap-independent translation initiation that depends on internal ribosomal entry sites.[26,39,40] It has been hypothesized that excess secondary structure including RG4s in 5′-UTR may have a regulatory function on translation.[26,27,37−40] Indeed, as observed with RNA hairpins, RG4s can either inhibit the assembly of the translation initiation machinery at the 5′-cap or interfere with the scanning of the ribosome toward the AUG codon (Supporting Information S1).[26,41−47] Luciferase data obtained with expression plasmids designed with G4 in specific positions upstream of AUG have been reported as a proof-of-principle.[26,43,44] In our work we demonstrate that furan- and thiophene-fused anthraquinone derivatives bind to RG4s in the 5′-UTR of *KRAS* mRNA under low-abundance cellular conditions. These G4-specific molecules are found to suppress the *KRAS* oncogene in pancreatic cancer cells through a mechanism involving the inhibition of translation.

**Human *KRAS* 5′-UTR Forms RNA G-Quadruplex Structures.** The human *KRAS* transcript contains a 5′-UTR of 192 nucleotides, characterized by a high GC content (77%) (Figure 1). Its putative secondary structure can be predicted by Mfold,[48] a bioinformatic tool that generates a secondary structure by taking into consideration the orthologues of the RNA sequence. However, Mfold does not take into account G4 motifs; thus, one wonders if the secondary structure predicted for the *KRAS* 5′-UTR is actually formed, considering that it


DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

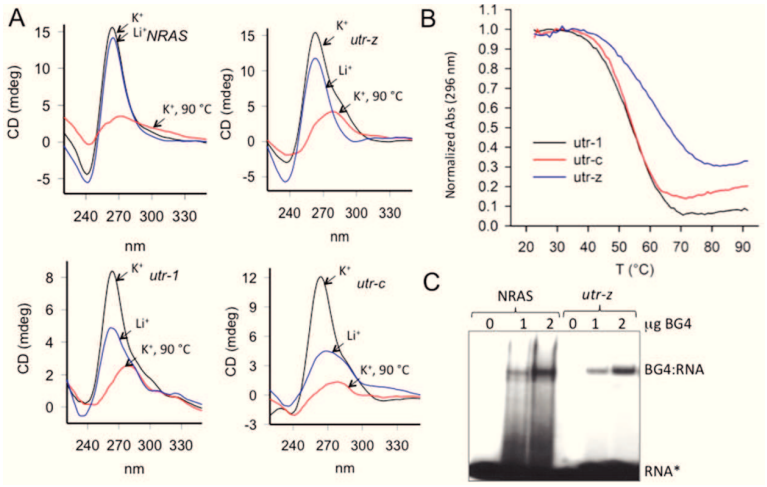



**Figure 2.** (A) Circular dichroism spectra at 25 °C and 90 °C of 5 $\mu$M *NRAS*, *utr*-1, *utr*-c, and *utr*-z RG4s in 50 mM Tris-HCl, pH 7.4, 100 mM KCl (or 100 mM LiCl); for *NRAS*, the KCl or LiCl concentration was 20 mM. The spectra were obtained in 0.5 cm cuvettes. (B) UV-melting curves of 5 $\mu$M *utr*-1, *utr*-z, and *utr*-c in 50 mM Na-cacodylate, pH 7.4, 100 mM KCl or LiCl. The curves were obtained by measuring the absorbance at 296 nm. The absorbance was normalized with the value at 20 °C. (C) EMSA in 50 mM Tris-HCl, pH 7.4, 100 mM KCl of mixtures containing *NRAS* RG4 and BG4 (lanes 2 and 3 from left) or *utr*-z RG4 and BG4 (lanes 5 and 6 from left). BG4 was used at 1 and 2 $\mu$g, and the gel was 5% polyacrylamide in TBE.

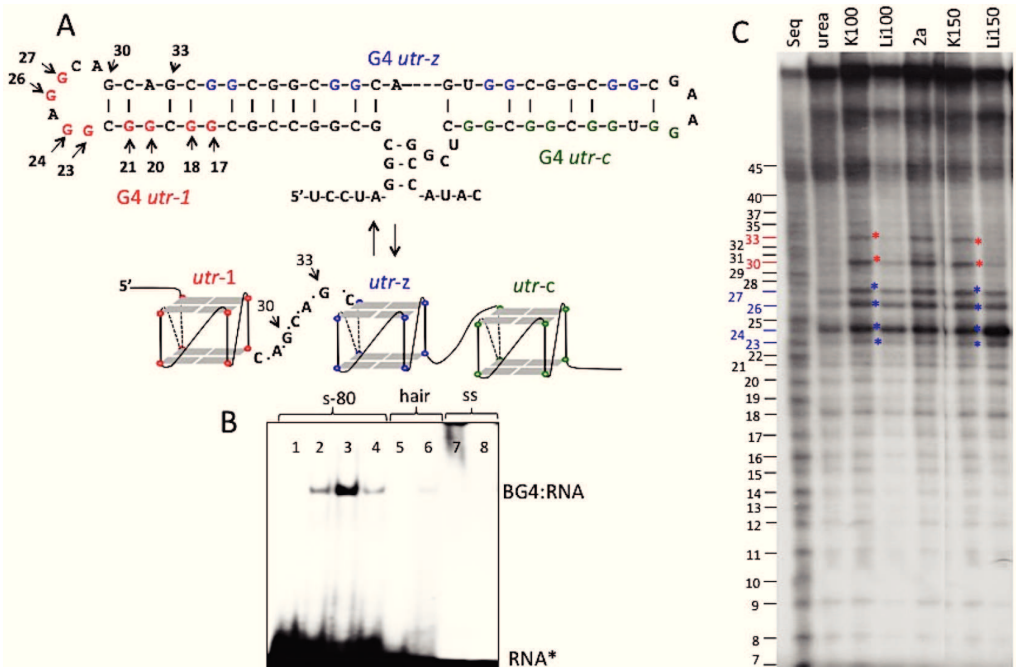



**Figure 3.** (A) Stem-loop structure of the first 80 nt of *KRAS* 5′-UTR (s-80) given by Mfold. Sequence s-80 is in equilibrium with an alternative structure characterized by three nonoverlapping G4 RNAs. (B) EMSA in 5% polyacrylamide, TB 1X, of mixtures between BG4 and s-80 or RNA hairpin (hair) 5′-GGCCGCCGCAGUGGCGGCGG or ssRNA 5′-UGUAAACAUCCUACACUCAGCU in 50 mM Tris-HCl, pH 7.4, 100 mM KCl ($K^+$-buffer) or 100 mM LiCl ($Li^+$-buffer). Except mixture 4, all have been prepared in $K^+$-buffer. Loading: s-80 (lane 1); s-80 + 1 $\mu$g BG4 (lane 2); s-80 + 2 $\mu$g BG4 (lane 3); s-80 + 1 $\mu$g BG4 in $Li^+$-buffer (lane 4); hair (lane 5), hair +1 $\mu$g BG4 (lane 6); ssRNA (lane 7); ssRNA + 1 $\mu$g BG4 (lane 8). (C) RNase T1 footprinting of s-80. Loading from left to right: alkaline RNA fragmentation (Seq); RNase T1 reaction in urea (urea), in $K^+$-buffer (K100), in $Li^+$-buffer (Li100), in $K^+$-buffer + **2a** ($r$ = 4) to see if the hairpin = G4 equilibrium is shifted by the ligand (**2a**), in $K^+$-buffer with 150 mM KCl (K150) in $Li^+$-buffer with 150 mM (Li150). The bottom band (band 7) matches the mobility of a 7-mer fragment.

contains 33 GG runs which give rise to a multitude of G4 motifs. To identify the G4 motifs with the highest G-score (i.e., highest propensity to fold into a G4), we interrogated QGRS Mapper.[49] We considered a consensus G4 motif composed by 2 G-tetrads and loop length up to 12 nt. The analysis gave three nonoverlapping G4 motifs with a G-score = 21 (Table 1). If the overlapping G4 motifs were included in the analysis, their number was >300, suggesting that the human *KRAS* 5′-UTR sequence has a high propensity to form RG4s. The three nonoverlapping G4 motifs are located within the first 80 nt of


DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

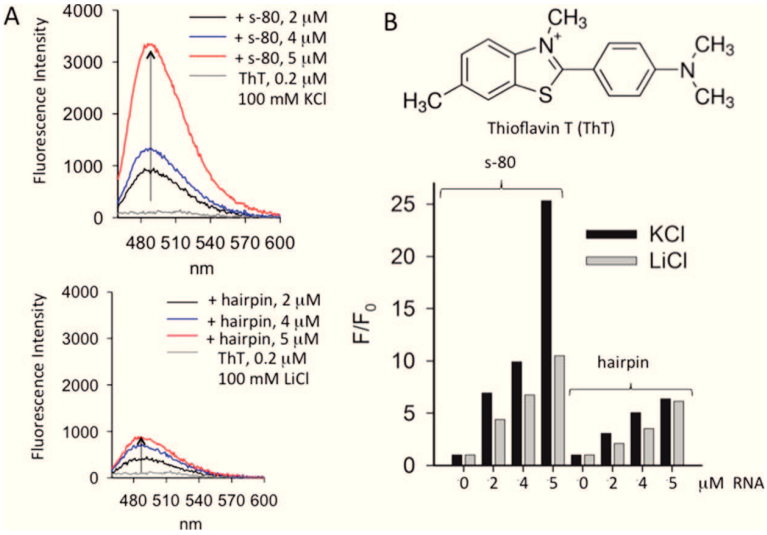



**Figure 4.** (A) Fluorescence emission spectra in 50 mM Tris-HCl pH 7.4, 100 mM KCl ($K^+$-buffer), of 0.2 $\mu$M ThT in the presence of increasing amounts of 80-mer RNA fragment s-80 (top panel) or RNA hairpin 5′-GGCCGCCGCAGUGGCGGCGG (bottom panel). (B) Structure of thioflavin T (ThT) and increase of fluorescence quantum yield at 485 nm of ThT following addition of increasing amounts of s-80 or hairpin in $K^+$-buffer and $Li^+$-buffer (50 mM Tris-HCl pH 7.4, 100 mM LiCl).

5′-UTR, and each displays a circular dichroism spectrum characterized by a strong ellipticity at 265 nm and a negative ellipticity at 240 nm, typical of parallel G-quadruplexes, as does the well-known G4-motif located in the 5′-UTR of *NRAS* (Figure 2 A).[26,50] The G4 motifs show UV-thermal difference spectra characterized by a negative peak at 295 nm in $K^+$, but not in $Li^+$, buffer, a distinctive feature of G4 structures (Supporting Information S2).[51] They show cooperative UV-melting curves at 296 nm (as well as at 260 nm), with $T_M$ varying from 53 °C to 64 °C in $K^+$ buffer (Figure 2 B and Table 1). The melting curves analyzed with a two-state model gave $\Delta G$ of quadruplex formation between ∼−3.7 and −5.6 kcal/mol (Table 1).

To confirm that the selected G4 motifs adopt a RG4 structure, we analyze by EMSA if they are recognized by BG4, an antibody specific for G-quadruplexes.[52,53] This analysis was performed with only the 20-mer *utr-z* G4 motif, as *utr-1* and *utr-c* are too short for binding to BG4. We found that BG4 formed a stable RNA−protein complex with *utr-z* as well as with the *NRAS* RG4, which was used as a positive control (Figure 2 C).[26] Next, we analyzed by EMSA the first 80-nt of 5′-UTR (s-80), which contains the three nonoverlapping G4 motifs. We wanted to find out if the G4 structures are also present in this longer RNA s-80 sequence, although it could form a mismatched stem-loop structure, according to Mfold (Figure 3 A). We found that in $K^+$-buffer, BG4 clearly bound to s-80, while it did not essentially bind to single-stranded RNA or hairpin RNA sequences, thus suggesting that *KRAS* 5′-UTR forms indeed local G4 structures (Figure 3 B). We found that BG4 bound to s-80, although with a weaker intensity, also in $Li^+$-buffer. This is probably due to the fact that s-80 exists as a hairpin = G4 equilibrium, which is more shifted to the right in $K^+$- than $Li^+$-buffer. BG4, upon binding to the existing quadruplex, is likely to push the equilibrium to the right. The capacity of RNA sequences to fold into G4s in the absence and presence of various metal ions, including $K^+$, $Na^+$, and $Li^+$, has been examined by Miserachs et al.[54] Further evidence that G4 is present in s-80 was obtained by measuring the reactivity of the guanines with RNase T1,[55] taking into account that the guanines involved in the formation of Watson−Crick or Hoogsteen hydrogen bonds do not react with RNase T1. If s-80 assumes the stem-loop structure of Figure 3 A, the loop guanines G23-G24 and G26-G27 should be reactive. On the other hand, if s-80 forms the three nonoverlapping G4 structures, the loop guanines should not be reactive, as they should be involved in the formation of the G-tetrads. In contrast, guanines G30 and G33, falling between the *utr-1* and *utr-z* RG4 structures, should be reactive to RNase T1. The footprinting of s-80 over the loop region shows that there is a prevalence of stem-loop structure in $Li^+$-buffer, as G30 and G33 are not or very slightly reactive (Figure 3 C, from left lanes 4 and 7). By contrast, in $K^+$-buffer both G4 and stem-loop structures are present in solution, as G30 and G33 as well as G23-G24 and G26-G27 are clearly reactive with RNase T1 (lanes 3 and 6 from left). This behavior suggests that s-80 exists in solution in equilibrium between G4 and the stem-loop structure. To further support this conclusion, we tested the presence of RG4s in s-80 with thioflavin T (ThT), a fluorescence sensor specific for G4.[56] As illustrated in Figure 4 A, ThT exhibited a strong increase of fluorescence emission upon binding to s-80 but not upon binding to a RNA hairpin (5′-GGCCGCCGCAGUGGCGGCGG). In Figure 4 B we compared the fluorescence enhancement of ThT induced by s-80 or hairpin RNA, in $K^+$- or $Li^+$-buffer.

$K^+$-buffer s-80 caused an increase of fluorescence ($F/F_0$) up to 25 while in $Li^+$-buffer only up to 10, in agreement with the fact that the hairpin = G4 equilibrium is more shifted to the right in the $K^+$-buffer. By comparison, a designed hairpin RNA increases $F/F_0$ up to 5. These data strongly suggest that under physiological conditions, RG4 structures are certainly present in the first portion of *KRAS* 5′-UTR. Recently, Weldon et al.[57] performed RNA footprintings on wild-type Bcl-x-681 transcript and its 7-deaza-G analogue, which cannot form RG4. They found that RG4 is present in the transcript, despite a possible formation of competing stem-loop structures. The Bcl-x G4


DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

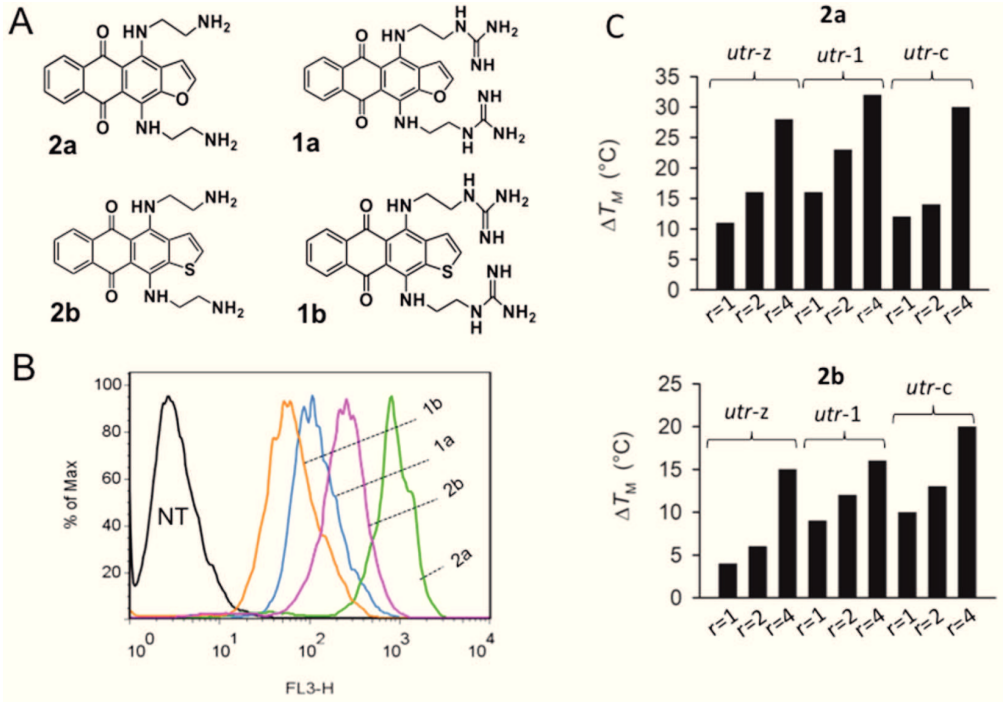



**Figure 5.** (A) Structures of ATFD **2a**, **1a** and ATPD **2b**, **1b**. Two amine- or guanidine-alkyl side chains have been attached to the ring system of the molecules. (B) Flow cytometry data showing the uptake of 3 $\mu$M compounds by Panc-1 cells after an incubation of 4 h. The two alkylamine compounds **2a** and **2b** are taken up more than the guanidine analogues. (C) $\Delta T_M$ is the $T_M$ increase of *utr*-z, *utr*-1, and *utr*-c G4 RNAs caused by **2a** and **2b** at $r$ = 1, 2, and 4, in 50 mM Tris-HCl, pH 7.4, 100 mM KCl.

motifs have G-scores similar or lower ($15 \leq$ G-score $\leq 21$) than those found in the 5′-UTR of *KRAS* (G-score = 21).

**ATPD and ATFD: Uptake and Capacity To Stabilize RG4.** As the 5′-UTR of *KRAS* is conserved in mammals (Supporting Information S3) and forms RG4s, we hypothesized that these unusual structures could be involved in the mechanism regulating translation of *KRAS*. Previous studies have showed that translation can be modulated by small molecules targeting to RG4s.[41,58,59] We therefore searched for molecules that obey the following criteria: (i) high capacity to penetrate cell membranes; (ii) binding to *KRAS* mRNA despite its typical cellular low-abundance; (iii) high affinity for RG4s. In previous studies we reported that anthrathiophenediones (ATPD) with alkyl side chains carrying either guanidine or chloroacetamidine terminal groups penetrate bladder cancer cells and bind to DNA and RNA G4s.[60,61] Because the side chains strongly impact the uptake of these molecules,[60] we tested in human pancreatic Panc-1 cancer cells several ATPD analogues and focused on 4,11-bis[2-guanidinoethylamino)-anthra[2,3-*b*]thiophene-5,10-dione (**1b**) and 4,11-bis(2-aminoethylamino)anthra[2,3-*b*]thiophene-5,10-dione (**2b**)[62] as well as on their furan analogues, anthrafurandiones (ATFD) **1a** and **2a**[63] (Figure 5 A). Both types of ligands have alkyl side chains ending either with guanidine or amine groups. As the chromophore of these molecules emits fluorescence upon excitation at 488 nm, we investigated their cellular uptake by FACS (Figure 5 B). The results showed that compounds **2a** and **2b** with aminoethyl side chains are taken up 20- and 4-fold more efficiently than the corresponding guanidine analogues **1a** and **1b**, respectively. An explanation can be that compounds carrying a localized charge generally display lower membrane permeability than neutral compounds. Compound **2a** differs from **2b** only for an atom in the five-member ring: oxygen against sulfur. Nonetheless, the former is 5-fold more permeable to Panc-1 cells than the latter. The higher polarizability of sulfur compared to oxygen provides a rationale for this behavior.[64] Due to their high uptake in Panc-1 cells, we used **2a** and **2b** to design a strategy aiming at inhibiting translation of *KRAS* in Panc-1 cells. The guanidine analogues **1a** and **1b** were used as reference compounds: because their uptake is lower, they are expected to produce a weaker cellular effect. We first asked if the molecules bind to and stabilize the RG4 structures. UV-melting experiments showed that **2a** and, to a lesser extent, **2b** strongly stabilize the three nonoverlapping RG4s located in the *KRAS* 5′-UTR. The increase of $T_M$ ($\Delta T_M$) of the three RG4s in the presence of the G4 ligands at $r$ = 1, 2, and 4 ($r$ = [ligand]/[G4]) are reported in the histograms of Figure 5 C. **2a** caused a $T_M$ increase up to 32 °C, while **2b** up to 20 °C. The lower $\Delta T_M$ brought about by **2b** is probably due to the polarizability of sulfur that may reduce the stacking of the chromophore upon the G-tetrads. Our data show that when thiophene is replaced with furan, the ligand increases not only its capacity to stabilize RG4 but also the uptake in Panc-1 cells. We also investigated the quadruplex-to-duplex specificity of the G4 ligands by competition experiments. For instance, the $T_M$ of *utr-1* RG4 in the presence of **2a** at $r$ = 4 did not change when a 5- or 10-fold excess of 21-mer RNA duplex was added to the mixtures (not shown). This is in keeping with our previous data on analogue ligands.[60,61]

Finally, the affinity of ligands **2a** and **2b** for various G4 RNAs located in 5′-UTR was determined by fluorescence titrations. A typical titration obtained with *utr*-1 and **2a** is shown in Supporting Information S4. The various $K_D$s are reported in Table 2. The average $K_D$ value for **2a** is ~140 nM and for **2b** is



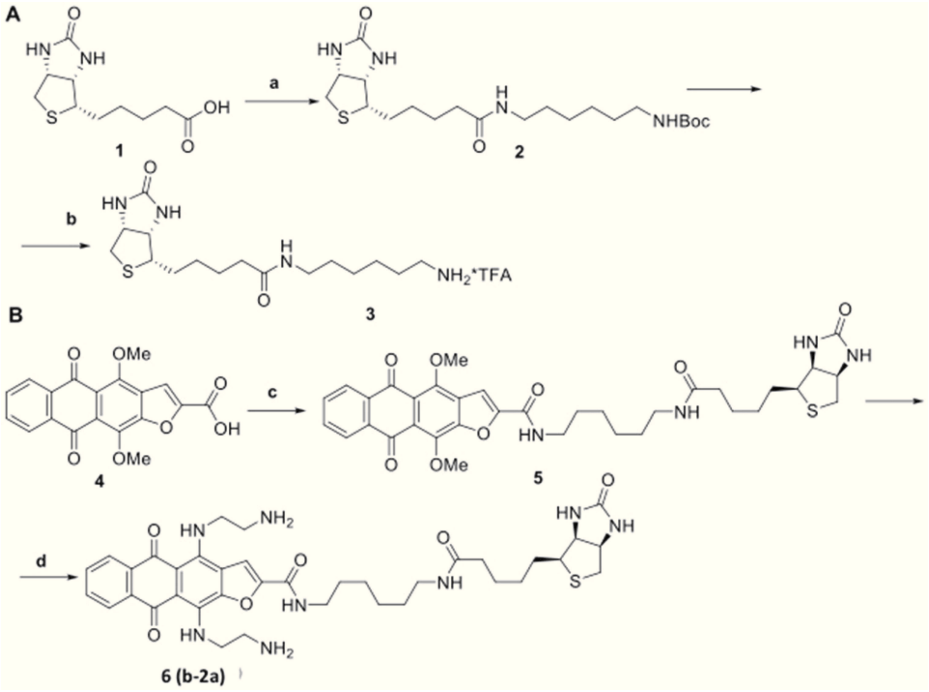
DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

Table 2. $K_D$ Values Relative to the Interaction between *KRAS* G4 RNAs and Ligands 2a and 2b

| | sequence (5′→3′) | 2a $K_D$ (nM)[a] | 2b $K_D$ (nM)[a] |
|---|---|---|---|
| *utr*-1 | GCGGCGGCGGAGGCA | 75 ± 6 | 91 ± 15 |
| *utr*-z | GGCGGCGGCAGUGGCGGCGG | 251 ± 151 | 294 ± 138 |
| *utr*-4[b] | CAGCAGCGGCGGCGGCAGUGG | 96 ± 45 | 83 ± 27 |
| WC hairpin[c] | CCGCCGCAGUGGCGGCGG | 514 ± 103 | 482 ± 109 |

[a]Obtained from fluorescence titrations in 50 mM Tris-HCl, pH 7.4, 100 mM KCl. [b]G4 motif present in the *KRAS* 5′UTR that overlaps other G4 motifs; [c]WC RNA hairpin with 7-C:G stem and 4-nt loop (AGUG) .

Scheme 1. Synthesis of Biotinylated Anthrafurandiones b-2a[a]




[a]Reagents and conditions: (a) BocNH$(CH_2)_6NH_2$, HATU, NMM, DMF, rt; overnight; yield 89%; (b) DCM, TFA, rt; 3 h; yield 87%; (c) **3**, PyBOP, DIPEA, DMSO, rt, 1 h, yield 80%; (d) ethylendiamine, THF, 50 °C, 1.5−2 h, yield 68%.

~156 nM. The $K_D$ for the binding of the ligands to competing mismatched hairpin RNAs could not be measured, as we were unable to find experimental conditions in which these putative structures, predicted by Mfold, are stable. We also designed a non-natural RNA stem-loop structure and found that **2a** and **2b** have affinities for this hairpin from 2 to 6 times lower than that for the RG4s. Interestingly, the $K_D$s for the binding of the ligands to the critical G4 DNA formed in the *KRAS* promoter are 626 ± 71 nM for **2a** and 278 ± 21 nM for **2b**. Although the ligands can bind also to G4 DNA, the fact that they accumulate more in the cytoplasm than in the nucleus (Supporting Information S5) and that there are many copies of mRNA RG4s, suggest that the main target, i.e., those producing a stronger cellular effect, is likely to be in mRNA.

**ATFDs Target the *KRAS* Transcript under Cellular Low-Abundance Conditions.** Studies so far reported on the use of small molecules to inhibit translation are based on luciferase assays and on the assumption that the G4 ligands bind to mRNA.[26,43−45] The first issue that we addressed in our study was to demonstrate that **2a** (the compound showing the highest uptake) does indeed bind to *KRAS* mRNA, under its low-abundance cellular condition. To do this, we synthesized biotinylated **2a** (b-**2a**) and set up a biotin−streptavidin pull-down assay.

*Synthesis of Biotinylated **2a** (b-**2a**).* The synthesis of compounds **1a**, **1b**, **2a**, and **2b**, based on the nucleophilic substitution of alkoxy groups in the peri-positions of heteroarene-fused anthracenediones, has been described previously.[62−65] To link biotin to **2a** we used as spacer 1,6-diaminohexane.[66−68] First, biotin (**1**) was coupled with mono-Boc-protected 1,6-diaminohexane using HATU as coupling reagent. The subsequent cleavage of the protecting group led to building block HDA−biotin (**3**) (Scheme 1 A).[66] As starting compound for the biotinylation of **2a**, we used 4,11-dimethoxy-5,10-dioxoanthra[2,3-*b*]furan-2-carboxylic acid (**4**).[69] The condensation of the amine group of **3** with the carboxylic group of **4** in the presence of PyBOP yielded the biotinylated 4,11-dimethoxy intermediate **5** (Scheme 1 B). Finally, by treating 4,11-dimethoxyanthra[2,3-*b*]furan-2-carboxamide (**5**) with ethylenediamine in THF at 50 °C we obtained to biotinylated 4,11-bis(aminoalkylamino)anthra[2,3-*b*]furan-5,10-dione **6** (b-**2a**).

*Streptavidin−Biotin Pull-Down Assays with b-**2a**.* To demonstrate that the ligands bind to RG4s in *KRAS* mRNA, we first ask if, by means of a biotin−streptavidin assay, b-**2a** is



DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

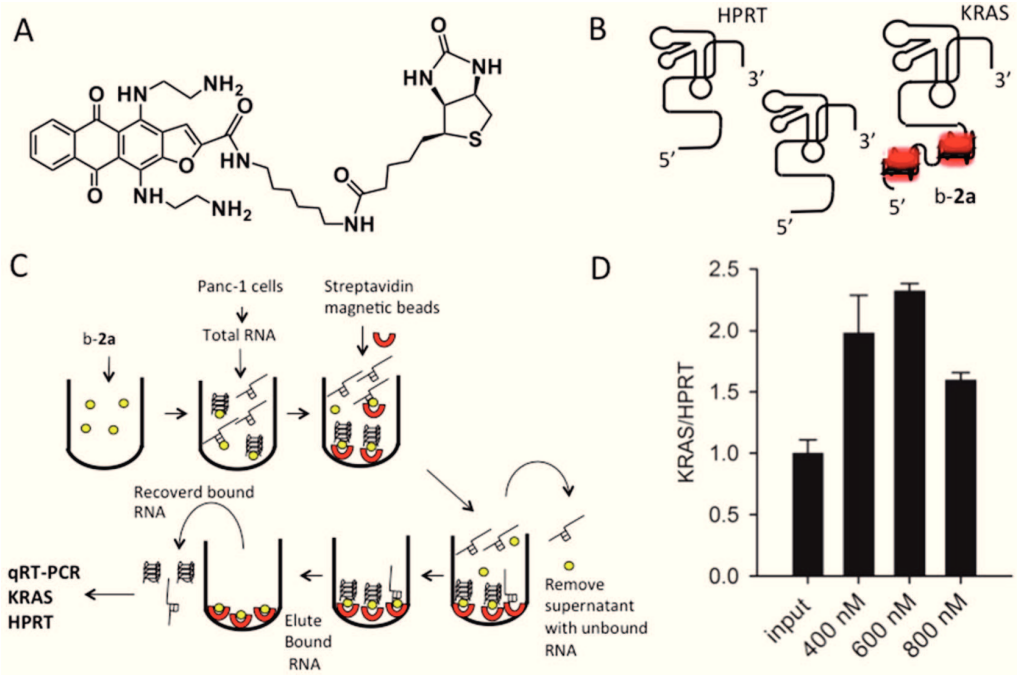



**Figure 6.** (A) Structure of b-**2a** conjugated with biotin. (B) Total cellular mRNAs contain *KRAS* mRNA with G4 RNA in the 5′-UTR. G4-RNA is bound by b-**2a**. *HPRT* is not bound or little bound by b-**2a**. (C) The streptavidin−biotin pull-down assay: b-**2a** pulls down preferentially G4-motif mRNAs. The recovered mRNA was used to determine by RT-qPCR the amounts of *KRAS* and *HPRT* mRNAs. (D) The histograms show the ratio of *KRAS*/*HPRT* mRNAs in the input, i.e., total cellular extract (fixed to 1), and in the recovered RNA from a total cellular extract treated with increasing amounts of b-**2a**. An enrichment of *KRAS* over *HPRT* of nearly 3 times was obtained, suggesting that the biotinylated ligands bind to *KRAS* mRNA within the total cellular extract.

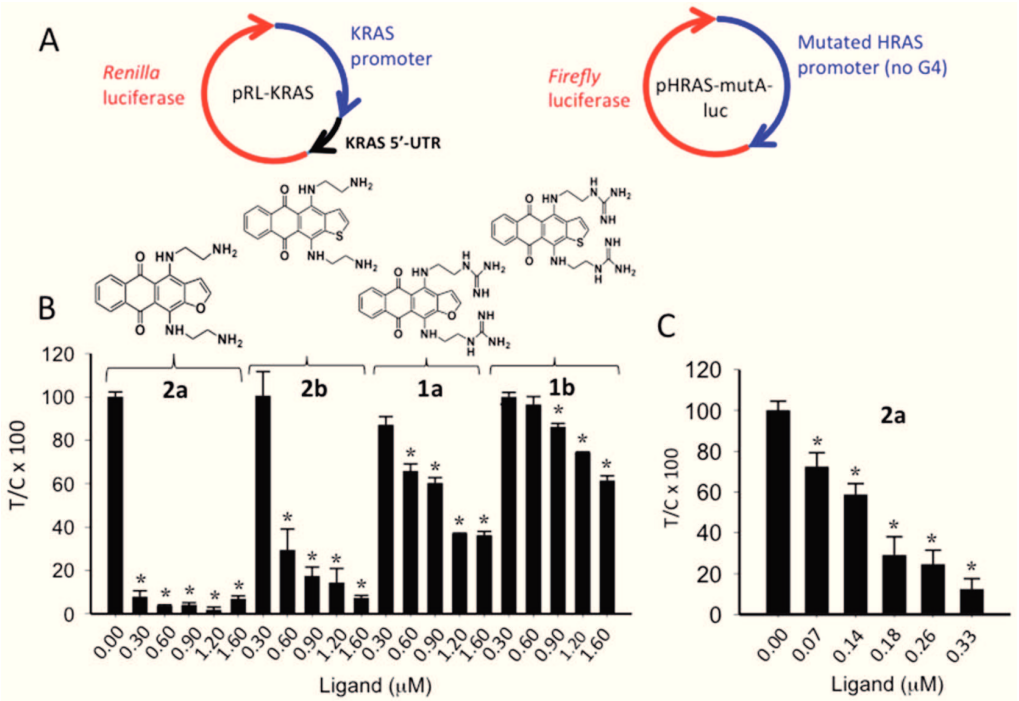



**Figure 7.** (A, B) Dual luciferase assay showing the effect of compounds **2a**, **2b**, **1a**, and **1b** (0−1.6 μM) on *Renilla* luciferase driven by the *KRAS* promoter comprising the 5′-UTR (pRL-KRAS). As reference, we used a plasmid in which *Firefly* luciferase was driven by the *HRAS* promoter mutated to abolish its capacity to form quadruplex structures.[61] The assay shows that **2a** strongly inhibits luciferase expression. Also **2b** shows a strong inhibitory activity. (C) When Panc-1 cells are treated with **2a** at lower concentrations (0−0.33 μM), a dose−response reduction of luciferase is observed. T = *Renilla*/*Firefly* in treated cells, C = *Renilla*/*Firefly* in untreated cells. *$P < 0.05$.



DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

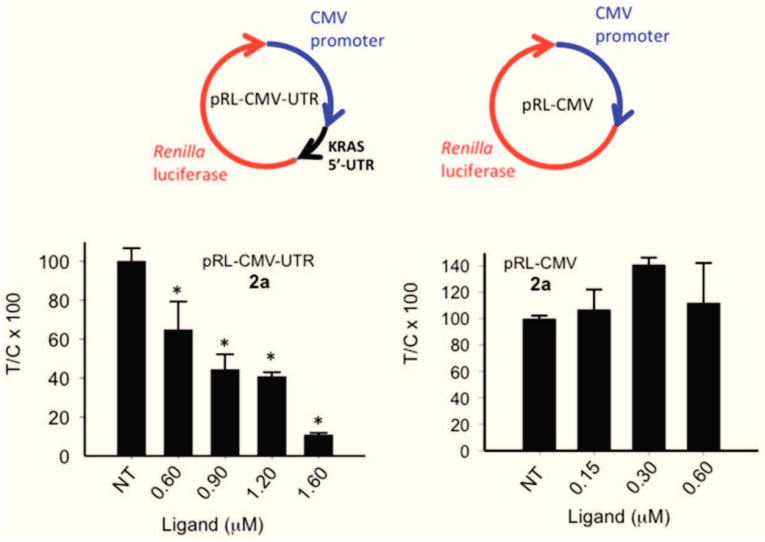



**Figure 8.** Dual luciferase assay with *Renilla* luciferase plasmid pRL-CMV-UTR or pRL-CMV and *Firefly* luciferase plasmid pHRAS-mutA-luc. Left histograms show a strong dose−response luciferase reduction caused by increasing amounts of **2a** (0−1.6 μM), while the right panel shows that when the *KRAS* 5′-UTR is removed, the luciferase reduction is not observed, suggesting that the inhibitory effect is mediated by 5′-UTR. T = *Renilla*/*Firefly* in treated cells, C = *Renilla*/*Firefly* in untreated cells. *$P < 0.05$.

capable to pull down a DNA containing a G4 structure (Figure 6 A). We designed two DNA strands of 115 and 89 nt of which the former contained in the middle a G4 motif (telomeric *htel*[70]). Both DNA strands had the same 5′- and 3′-ends and could be amplified with the same couple of primers. A 1:1 mixture of the two strands was incubated with an excess b-**2a**, and the DNA bound to the biotinylated ligand was pulled down by streptavidin−magnetic beads and the recovered DNA amplified. On a calibration curve, the DNA recovered was dramatically enriched with the 115-nt sequence containing G4, suggesting that b-**2a** efficiently pulled down the DNA strand carrying G4: a prerequisite essential for the next step of the experiment (Supporting Information S6).

The pull-down experiment was repeated by replacing the 1:1 mixture with total cellular RNA extracted from Panc-1 cells. Cellular RNA is composed by ribosomal, transfer, and, in minimal part, mRNA. As illustrated in Figure 6 B, the transcriptome contains mRNAs without G4 structures (such as *HPRT*) and with G4 structures in 5′-UTR as *KRAS*. We reasoned that while all transcripts will have weak binding sites for b-**2a** (RNA stem-loop secondary structures), only a fraction of transcripts will exhibit both weak (stem-loop) and strong (G4) binding sites for b-**2a**. On the basis of this assumption, total cellular RNA from Panc-1 cells was incubated with b-**2a** and the RNA bound to the ligand was pulled down with streptavidin-coated magnetic beads. The amounts of *KRAS* and *HPRT* transcripts in the recovered RNA were determined by quantitative RT-PCR and compared to the amounts of the same genes detected in the input (untreated cellular RNA). The results showed that *KRAS*/*HPRT* in the pulled-down RNA increased nearly 3-fold compared to the input, suggesting that b-**2a** binds indeed to RG4s in *KRAS* transcripts (Figure 6 C,D). The pull-down experiment was repeated with increasing concentrations of b-**2a** (from 400 to 800 nM). The highest enrichment in *KRAS* transcript was obtained at a concentration of **2a** of 600 nM. With higher concentrations of b-**2a**, the enrichment decreased as the ligand bound to both strong (G4) and weak (stem-loop) sites (Figure 6 B). Together, the data demonstrate that a small molecule such as **2a**, designed to inhibit gene expression, targets G4 RNA in the 5′-UTR of the *KRAS* transcript under cellular conditions in which the transcript is typically in low abundance.

**Capacity of Compounds 2a and 2b To Repress Translation in Pancreatic Cancer Cells.** After having demonstrated that ATPD and ATFD bind to RG4s in the *KRAS* transcript, we asked if they inhibit the translation of the oncogene in pancreatic cancer cells. First, we performed a dual-luciferase assay as a proof-of-principle. We used a plasmid where *Renilla* luciferase was driven by the *KRAS* promoter including 5′-UTR (pRL-KRAS). To determine the transfection efficiency, we used pHRAS-mutA-luc, in which *Firefly* luciferase is driven by a mutated *HRAS* promoter, which does not bear G4 motifs (Figure 7 A).[71] Panc-1 cells were first treated for 6 h with increasing amounts of **2a** and **2b**, with the guanidine analogues **1a** and **1b** (from 0.3 to 1.6 μM), and then with the luciferase vectors. The ratio between *Renilla* and *Firefly* luciferases was measured 48 h after transfection. **2a** and **2b** at concentrations between 0.3 and 1.6 μM dramatically lower luciferase, much more than that of the reference guanidine analogues (Figure 7 B). When **2a** was used at lower concentrations, between 0.07 and 0.33 μM, a clear dose−response inhibitory effect was observed (Figure 7 C). If one considers that the four ligands have similar affinity for *utr-z* (and also for the other RG4s) but different uptake in Panc-1 cells, there is a clear correlation between luciferase inhibition and uptake. The higher the uptake (**2a** and **2b**), the higher the luciferase inhibition.

It might be argued that the luciferase repression mediated by **2a** and **2b** is due to the binding of the ligands not only to the RG4s in 5′-UTR but also to G4 in the *KRAS* promoter.[23,72,73] To address this point, we prepared a vector where *Renilla* luciferase is driven by the cytomegalovirus (CMV) promoter. At the 3′-end of CMV promoter, we inserted the *KRAS* 5′-UTR (pRL-CMV-UTR). As the CMV promoter does not form G4 structures, the strong binding sites for **2a** are only located in the 5′-UTR sequence. As a reference vector, we used pHRAS-


DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

mutA-luc. The dual luciferase assay showed that **2a** represses *Renilla* luciferase expression in a dose−response manner, suggesting that the presence of *KRAS* 5′-UTR alone is sufficient to promote the repression of luciferase (Figure 8, left panel). The inhibition is observed with higher ligand concentrations than those used with plasmid pRL-KRAS, because the CMV promoter is stronger than the *KRAS* promoter. We then measured the effect of **2a** on pRL-CMV, which lacks the *KRAS* 5′-UTR downstream of the CMV promoter. In this case **2a** did not repress *Renilla* luciferase, in agreement with the fact that G4 is not present in this construct (Figure 8, right panel). This experiment supports the conclusion that **2a** upon binding to RG4 in the 5′-UTR of *KRAS* inhibits translation. We cannot rule out that the compounds also bind to the *KRAS* promoter. However, considering that (i) the G4 ligands accumulate more in the cytoplasm than in the nucleus (Supporting Information S5) and that (ii) there are many copies of RNA targets in the cells, even under mRNA low-abundance cellular conditions, **2a** and **2b** are expected to bind more favorably to G4 in mRNA than in promoter DNA.

**Effect of 2a and 2b on *KRAS* Expression in Pancreatic Cancer Cells.** As the designed compounds, in particular **2a**, decrease the luciferase from vectors containing the *KRAS* 5′-UTR, we asked if these G4 ligands are also able to inhibit the expression of genomic *KRAS* in pancreatic cancer cells. We measured by quantitative RT-PCR the level of *KRAS* transcript in Panc-1 cells after 6 and 24 h of treatment. The results showed that while **2b**, **1a**, and **1b** did not lower the level of mRNA, **2a** reduced it to about 50% of the control, at both time points. By contrast, all the compounds displayed a strong capacity to suppress the KRAS protein. In particular **2a** and **2b**, which are efficiently taken up by Panc-1 cells, brought the protein down to <10% of the control (Figure 9 A,B). In light of these results, we concluded that **2a** and **2b** inhibit *KRAS* mainly at the translational level. Moreover, both luciferase and Western blot data show that compound **2a** is slightly more active than **2b**, possibly because **2a** has a higher capacity to penetrate the cell membrane than **2b**, and a fraction of it may also target G4 in the *KRAS* promoter.

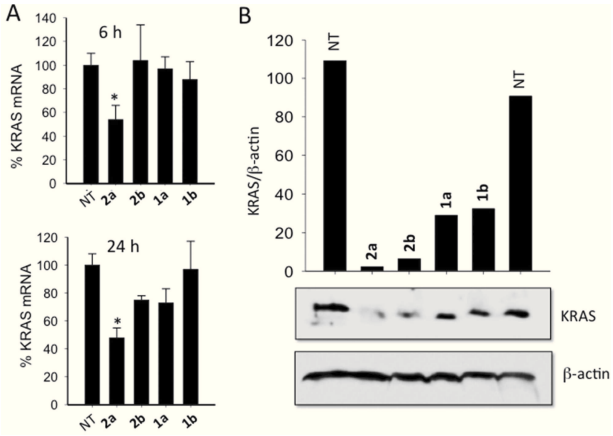



**Figure 9.** (A) Quantitative RT-PCR of *KRAS* mRNA in Panc-1 cells treated for 6 and 24 h with 1.6 $\mu$M **2a**, **2b**, **1a**, and **1b**. Ordinate reports the level of KRAS mRNA normalized to $\beta$2-microglobulin and *HPRT*. (B) Western blot determination of *KRAS* protein and $\beta$-actin in Panc-1 cells treated with 1.6 $\mu$M **2a**, **2b**, **1a**, and **1b** for 48 h. *$P$ < 0.05.

Pancreatic cancer cells, being addicted to *KRAS*, respond to the repression of the oncogene by activating apoptosis.[74] We found indeed that **2a** and to a lesser extent **2b** caused a strong activation of caspases 3/7 (Figure 10 A). In contrast, the reference compounds **1a** and **1b** showed a much weaker caspase activation. To confirm the activation of apoptosis, we also performed an annexin/propidium iodide assay (Figure 10 B). This is based on the observation that an early event occurring in apoptosis is the translocation of phosphatidylserine from the inner to the outer leaflet of the plasma membrane, thus exposing the phospholipid to the external cell environment. Annexin V, a phosphatidylserine recognizing protein labeled with FITC, can be used to detect this event by FACS. Early and late apoptosis can be distinguished by using both annexin V and propidium iodide (PI).[75] The experiment showed that 1.6 $\mu$M **2a** and **2b** increased the population of apoptotic cells 72 h after treatment: untreated Panc-1 cells (apoptotic cells ∼1.2%); Panc-1 cells treated with **2a** (apoptotic cells ∼32.5%); Panc-1 cells treated with **2b** (apoptotic cells ∼20.6%).

Finally, the metabolic activity of Panc-1 cells treated with the designed compounds was evaluated by a resazurin assay. Compound **2a** causes an inhibition of the metabolic activity ($IC_{50}$ = 0.26 $\mu$M) stronger than **2b** ($IC_{50}$ = 0.9 $\mu$M) (Figure 10 C). Moreover, **2a** caused a strong inhibition (∼80%) of colony formation in two pancreatic cancer cells: Panc-1 and BxPC-3 (Supporting Information S7).

## CONCLUSION

By setting up a streptavidin−biotin pull-down assay, we have demonstrated that **2a**, an anthrafurandione with amino-ethylamino side chains, which efficiently internalizes and accumulates in the cytoplasm of Panc-1 cancer cells, targets *KRAS* mRNA under low-abundance cellular conditions. Luciferase assays with specific vectors showed that **2a** and its anthrathiophenedione analogue **2b** repress translation in a dose−response manner, suggesting that they have a great potential in cancer therapy. Indeed, Western blots showed that these molecules strongly decrease the *KRAS* protein in Panc-1 cancer cells. Moreover, the compounds activate apoptosis, as indicated by the caspase 3/7 and annexin/propidium iodide assays, and reduce the metabolic activity as well as the colony formation of the cells.

The mechanism of action of the designed compounds **2a** and **2b** is based on their capacity to bind to G4 structures located in the 5′-UTR of *KRAS* mRNA. The presence of these folded structures in mRNA has been demonstrated by immunostaining in fixed cells.[76] A recent study by Guo and Bartel casts doubts on the existence of RG4s under in vivo conditions, as the authors found that RG4s are globally unfolded by single-stranded binding proteins.[77] If this is also true for the *KRAS* transcripts in pancreatic cancer cells, compounds **2a** and **2b** could inhibit translation by competing with the binding of single stranded-binding proteins to the RNA G4-motifs.

Finally, having established in cancer cells that **2a** and **2b** have a potential as anticancer agents, the next step will be in vivo testing. As pancreatic cancer cells are addicted to *KRAS*, therapeutics targeting this oncogene should be much more injurious for the malignant cells than for normal cells.



DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

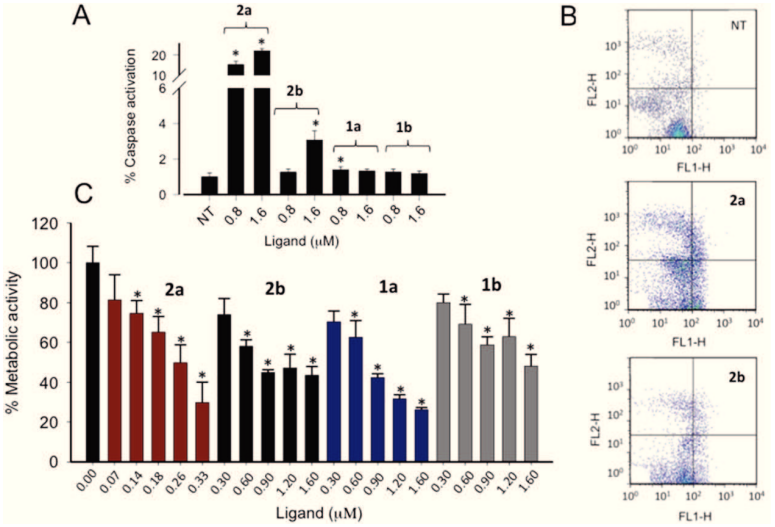



**Figure 10.** (A) Activation of caspase 3/7 in Panc-1 cells treated with **2a**, **2b**, **1a**, and **1b** (0.8 and 1.6 $\mu$M). Compounds **2a** and **2b** show activity higher than that of analogues **1a** and **1b**. (B) Annexin−propidium assay of Panc-1 cells treated for 72 h with **2a** (1.6 $\mu$M) and **2b** (1.6 $\mu$M). (C) Metabolic activity in Panc-1 cells treated with increasing concentrations of **2a**, **2b**, **1a**. and **1b**. Compound **2a** shows the highest activity. *$P < 0.05$.

## EXPERIMENTAL SECTION

**Oligonucleotides.** The oligonucleotides used in this study have been purchased from Microsynth (Switzerland). Oligonucleotide solutions in DEPC-treated Milli-Q water have been conserved at −80 °C. The sequences are reported in Table 1 and Supporting Information S8.

**Synthesis of 2a, 2b, 1a, 1b, and Biotinylated Ligand b-2a.** *General information.* NMR spectra were recorded on a Varian VXR-400 instrument operated at 400 MHz ($^1$H NMR) and 100 MHz ($^{13}$C NMR). Chemical shifts were measured in DMSO-$d_6$ using tetramethylsilane as an internal standard. Analytical TLC was performed using silica gel $F_{254}$ plates (Merck) and column chromatography with SilicaGel Merck 60. Melting points were determined using a Buchi SMP-20 apparatus and are uncorrected. High-resolution mass spectra were recorded with electron-spray ionization using a Bruker Daltonics microOTOF-QII instrument. UV spectra were recorded on a Hitachi-U2000 spectrophotometer. HPLC was performed using a Shimadzu Class-VP V6.12SP1 system. All solutions were dried over $Na_2SO_4$ and evaporated at reduced pressure using a Buchi-R200 rotary evaporator at a temperature below 45 °C. All products were vacuum-dried at room temperature. All solvents, chemicals, and reagents were obtained commercially and used without purification. The ligands **2a**, **2b**, **1a**, and **1b** have been synthesized as previously described.[60,61] The purity of final conjugates b-**2a** was >95% as determined by HPLC analysis (Supporting Information S9).

*Synthetic Procedures. tert-Butyl (6-(5-((3aS,4S,6aR)-2-Oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)hexyl) carbamate (**2**, N-Boc-HDA−Biotin).* *N*-Boc-1,6-Diaminohexane (0.50 g, 2.05 mmol) was added to a stirred solution of biotin (**1**, 0.50 g, 2.05 mmol), *N*-methylmorpholine (NMM, 0.23 mL, 2.05 mmol) and 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU, 0.77 g, 2.05 mmol) in dry DMF (20 mL). The reaction mixture was stirred overnight at room temperature and diluted with water, and the product was extracted with ethyl acetate (2 × 25 mL). The extract was washed twice with water, dried, and evaporated. The residue was purified by using column chromatography with chloroform−methanol (1:0 → 3:1) as the eluting solvent. The solid precipitate was recrystallized from *n*-propanol to give 0.84 g (89%) of the white powder of *N*-Boc-HDA−biotin (**2**);[63] mp 173−176 °C; $^1$H NMR (400 MHz, DMSO-$d_6$) $\delta$ 7.73 (t, 1H, $J$ = 5.4 Hz, NH); 6.77 (t, 1H, $J$ = 5.0 Hz, NH); 6.43 (br s, 1H, NH-biotin); 6.36 (br s, 1H, NH-biotin); 4.31−4.28 (m, 1H, CH); 4.13−4.11 (m, 1H, CH); 3.11−3.06 (m, 1H, SC*H*$CH_2$); 3.01 (dd, 2H, $^1J$ = 6.2, $^2J$ = 7.8 Hz, $NCH_2$); 2.89 (dd, 1H, $^1J$ = 6.0, $^2J$ = 7.2 Hz, $NCH_2$); 2.81 (dd, 1H, $^1J$ = 5.2, $^2J$ = 7.5 Hz, SCH*H*); 2.57 (d, 1H, $J$ = 12.4 Hz, SC*H*H); 2.03 (t, 2H, $J$ = 7.5 Hz, $COCH_2CH_2$); 1.60−1.20 (m, 23H, $C(CH_3)_3$, $7CH_2$); HRMS (ESI) calculated for $C_{21}H_{39}N_4O_4S^+$ $[M + H]^+$ 443.2687, found 443.2673.

*N-(6-Aminohexyl)-5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno-[3,4-d]imidazol-4-yl)pentanamide Trifluoroacetate (HDA−Biotin, **3**).* A solution of *N*-Boc-HDA−biotin (**2**, 0.80 g, 1.81 mmol) in a mixture of DCM (10.0 mL) and TFA (2.0 mL) was stirred for 3 h at the room temperature. The solvent was evaporated and residue reprecipitated from warm water with acetone. The precipitated crystals were filtered, washed with acetone, and dried to yield 0.72 g (87%) of HDA−biotin (**3**); mp 92−95 °C; $^1$H NMR (400 MHz, DMSO-$d_6$) $\delta$ 7.76 (t, 1H, $J$ = 5.4 Hz, NH); 7.70 (br s, 3H, $NH_3$); 6.42 (br s, 1H, NH-biotin); 6.37 (br s, 1H, NH-biotin); 4.32−4.29 (m, 1H, CH); 4.14−4.11 (m, 1H, CH); 3.11−3.06 (m, 1H, SC*H*$CH_2$); 3.01 (dd, 2H, $^1J$ = 6.0, $^2J$ = 6.8 Hz, $NCH_2$); 2.82 (dd, 1H, $^1J$ = 5.0, $^2J$ = 7.5 Hz, SCH*H*); 2.76 (dd, 1H, $^1J$ = 6.0, $^2J$ = 7.2 Hz, $NCH_2$); 2.58 (d, 1H, $J$ = 12.6 Hz, SC*H*H); 2.04 (t, 2H, $J$ = 7.5 Hz, $COCH_2CH_2$); 1.64−1.24 (m, 14H, $7CH_2$); HRMS (ESI) calculated for $C_{16}H_{31}N_4O_2S^+$ $[M + H]^+$ 343.2162, found 343.2156.

*Biotinyl-N-(6-aminohexyl)-4,11-dimethoxy-5,10-dioxo-5,10-dihydroanthra[2,3-b]furan-2-carboxamide (**5**).* A mixture of 4,11-dimethoxyanthra[2,3-*b*]furan-5,10-dione-2-carboxylic acid (**4**;[66] 0.30 g, 0.85 mmol), ethyldiisopropylamine (DIPEA, 0.5 mL, 3.00 mmol), biotinyl-*N*-(6-aminohexyl)amine trifluoroacetate (**3**, 0.39 g, 0.85 mmol), and benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP, 0.52 g, 1.00 mmol) in DMSO (15 mL) was stirred at room temperature for 1 h. The reaction mixture was diluted with water, and the product was extracted with ethyl acetate (2 × 20 mL). The extract was washed twice with water, dried, and evaporated. The residue was purified by using column chromatography with chloroform−methanol (1:0 → 3:1) as the eluting solvent. The yield of the orange solid of **5** was 0.46 g (80%); mp 132−134 °C; $^1$H NMR (400 MHz, DMSO-$d_6$) $\delta$ 8.87 (t, 1H, $J$ = 5.9 Hz, NH); 8.05−8.03 (m, 2H, 6,9-H); 7.88 (s, 1H, 3-H); 7.82−7.80 (m, 2H, 7,8-H); 7.76 (t, 1H, $J$ = 5.3 Hz, NH); 6.44 (br s, 1H, NH-biotin); 6.37 (br s, 1H, NH-biotin); 4.31−4.27 (m, 1H, CH); 4.12 (s, 3H, OMe); 4.07 (s, 3H, OMe); 4.13−4.12−4.10 (m, 1H, CH); 3.28 (dd, 2H, $^1J$ = 6.2, $^2J$ = 6.8 Hz, $NCH_2$); 3.10−3.06 (m, 1H, SC*H*$CH_2$); 3.02 (dd, 2H, $^1J$ = 6.0, $^2J$ = 7.0 Hz, $NCH_2$); 2.81 (dd, 1H, $^1J$ = 4.9, $^2J$ = 7.5 Hz, SCH*H*); 2.56 (d, 1H, $J$ = 12.3 Hz, SC*H*H); 2.04 (t, 2H, $J$ = 7.1 Hz, $COCH_2CH_2$); 1.59−1.25 (m, 14H, $7CH_2$); $^{13}$C NMR (100 MHz, DMSO-$d_6$) $\delta$ 182.05 (C═O); 181.65 (C═O); 171.87 (N−C═O); 162.75 (N−CO−N);



DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

156.92 (N−C═O); 151.61 (C); 150.47 (2C); 142.60 (C); 133.99 (C); 133.80 (C); 133.64 (C); 127.30 (C); 123.33 (C); 120.43 (C); 133.98 (CH); 126.10 (CH); 126.00 (CH); 108.44 (CH); 61.81 (CH); 61.71 (CH); 61.07 ($OCH_3$); 59.22 ($OCH_3$); 55.47 (CH); 39.86 ($CH_2$); 38.95 ($CH_2$); 38.31 (CH2); 35.25 ($CH_2$); 29.15 ($CH_2$); 29.00 ($CH_2$); 28.24 ($CH_2$); 28.06 ($CH_2$); 26.19 ($CH_2$); 26.14 ($CH_2$); 25.38 ($CH_2$). HRMS (ESI) calculated for $C_{35}H_{41}N_4O_8S^+$ $[M + H]^+$ 677.2640, found 677.2612.

*4,11-Bis((2-aminoethyl)amino)-biotinyl-N-(6-aminohexyl)-5,10-dioxo-5,10-dihydroanthra[2,3-b]furan-2-carboxamide (**6**, b-**2a**).* A mixture of compound **5** (0.27 g, 0.40 mmol) and ethylenediamine (1.5 mL) in THF (5.0 mL) was heated at 50 °C for 2−3 h. During this time, the yellow color of the reaction mixture changed to dark blue, and after the complete conversion of **5** (as determined by TLC) the solution was cooled and quenched with water. An aqueous solution of HCl (1%) was added to make the pH = 8.0, the solution was saturated with NaCl, and the product was extracted with warm *n*-butanol (3 × 25 mL). The extract was washed twice with brine, dried, and evaporated. The residue was purified by column chromatography using chloroform−methanol−concentrated $NH_4OH$ (10:2:0 → 10:4:1) as the eluting solvent. The purified residue was dissolved in a warm aqueous solution of HCl (1 N) and reprecipitated with acetone. The precipitated crystals were filtered, washed with acetone, and dried to yield 0.22 g (68%) of the dark blue powder of dihydrochloride **6**; mp 204−206 °C (decomp); HPLC Kromasil-100-5-mkm C-18 column (4 × 250 mm, LW = 260 nm), eluent: A, $H_3PO_4$ (0.01 M), B, MeCN; gradient B 20 → 60% (30 min), elution time 9.4 min, purity 96%. $^1$H NMR (400 MHz, DMSO-$d_6$) $\delta$ 12.18 (t, 1H, $J$ = 5.0 Hz, NH); 11.36 (t, 1H, $J$ = 5.5 Hz, NH); 9.32 (t, 1H, $J$ = 5.4 Hz, NH); 8.55 (s, 1H,3-H); 8.41 (br s, 3H, $NH_3$); 8.25−8.22 (m, 2H, 6,9-H); 8.18 (br s, 3H, $NH_3$); 7.82 (br s, 1H, NH); 7.80−7.78 (m, 2H, 7,8-H); 6.44 (br s, 1H, NH-biotin); 6.39 (br s, 1H, NH-biotin); 4.31−4.28 (m, 1H, CH); 4.13−4.10 (m, 3H, CH, $NCH_2$); 4.08−4.06 (m, 2H, $2NCH_2$); 3.29−3.26 (m, 4H, $2NCH_2$); 3.10−3.06 (m, 1H, S*C*H$CH_2$); 3.05−3.00 (m, 2H, $NCH_2$); 2.81 (dd, 1H, $^1J$ = 5.1, $^2J$ = 7.3 Hz, SCH*H*); 2.57 (d, 1H, $J$ = 12.4 Hz, SC*H*H); 2.05 (t, 2H, $J$ = 7.3 Hz, CO$CH_2$$CH_2$); 1.60−1.30 (m, 14H, $7CH_2$). HRMS (ESI) calculated for $C_{37}H_{49}N_8O_6S^+$ $[M + H]^+$ 733.3490, found 733.3494.

**Cell Culture, Metabolic Activity, and Proliferation Assays.** Panc-1 and BxPC-3 cells (human pancreatic cancer cells) were maintained in exponential growth in Dulbecco's Modified Eagle's Medium (DMEM) containing 100 U/mL penicillin, 100 mg/mL streptomycin, 20 mM L-glutamine, and 10% fetal bovine serum (Euroclone, Italy). The cell lines have been genotyped by Microsynth (Switzerland) to verify their identity. They matched 100% with the DNA-profiles of Panc-1 (ATCC CRL-1469TM) and BxPC-3 (ATCC CRL-1687TM). The metabolic activity assay was performed by seeding the cells ($10 \times 10^3$ cells/well) in a 96-well plate. After 1 day, the cells have been treated with the compounds and after an incubation of 72 h the resazurin assay was performed following a standard procedure.

Colony forming assays have been carried out with Panc-1 and BxPC-3 cells plated on a 60 mm plate and treated with 0.25 and 0.5 $\mu$M **2a** or 1 and 1.5 $\mu$M **2b**. After 18 days, the cells were fixed and stained for 10 min with 2.5% methylene blue in 50% ethanol. Colonies with >50 cells were counted.

**Circular Dichroism Spectra and UV-Melting Curves.** CD spectra have been obtained on a JASCO J-600 spectropolarimeter, equipped with a thermostated cell holder, with 5 $\mu$M oligonucleotide solutions in 50 mM Na-cacodylate pH 7.4, 100 mM KCl or LiCl (RNase free). The CD spectra (for *KRAS utr*-1, *utr*-c, and *utr*-z) and *NRAS* were registered at 25 °C and 90 °C. Spectra were recorded in 0.5 cm quartz cuvette. The spectra were calculated with J-700 Standard Analysis software (Japan Spectroscopic Co, Ltd.). Each spectrum was recorded three times, smoothed, and the baseline subtracted.

UV melting curves were obtained by using JASCO V-750 UV−visible spectrophotometer equipped with a temperature control system that heats/cools the sample through Peltier technology (ETCS-761) (Jasco, Portland, OR). Melting curves were recorded at 260 and 296 nm in a 0.5 cm path length quartz cuvette, heating the sample from 20 °C to 100 °C. The samples were prepared at a final concentration of 5 $\mu$M in 100 mM KCl and 50 mM Na-cacodylate, pH 7.4. Incubation with increasing amounts of the molecules (**2a** and **2b**) ($r$ = 1, 2, and 4) was performed for 1 h at room temperature.

**RNase T1 Footprinting and Electrophoretic Mobility-Shift Assay.** Single-stranded RNAs were purified by PAGE and 5′-end labeled with T4 polynucleotide kinase (ThermoFisher, Waltham, MA) and [$^{32}$P]-ATP (PerkinElmer, Waltham, MA) for 1.5 h at 37 °C. RNase T1 footprinting was performed with 30 nM s-80 heated for 5 min at 85 °C and then incubated overnight at 25 °C in 10x Structure Buffer (RNase T1 Biochemistry grade, ThermoFisher) with 100 mM KCl or LiCl. The reactions were performed with 0.05 units of RNase T1 (ThermoFisher) for 10 min at 25 °C and stopped with 20 $\mu$L inactivation/precipitation buffer (ThermoFischer). RNA was precipitated at −80 °C for 2 h and centrifuged for 30 min at 13000 rpm. Precipitated RNA was resuspended with loading buffer (ThermoFisher), heated for 5 min at 95 °C, and electrophoresed on a 20% denaturing gel, pre-equilibrated at 55 °C in a Sequi-Gen GT Nucleic Acids Electrophoresis Apparatus (Bio-Rad, Hercules, CA), equipped with a thermocouple that allows precise temperature control. EMSA assays were performed with 20 nM *utr-z* or s-80 labeled at the 5′-end with [$^{32}$P]-ATP. The mixture was incubated for 30 min at 37 °C with increasing concentrations of antibody BG4 (1 and 2 $\mu$g). BG4 was produced according to Studier et al.[78] The samples were run in a 5% TBE 1x gel for 2 h. After running, the gel was fixed in a solution containing 10% acetic acid and 10% methanol, dried at 80 °C, and exposed to Hyperfilm MP (GE Healthcare) for autoradiography.

**Western Blots.** Total protein lysates (15 $\mu$g) were electrophoresed on 12% SDS-PAGE and transferred to a nitrocellulose membrane at 70 V for 2 h. The filter was blocked for 1 h with 5% BSA solution in PBS 0.05% Tween (Sigma-Aldrich, St. Louis, MO) at room temperature. The primary antibodies used are antiactin (clone JLA20, IgM mouse, $1 \times 10^{-4}$ $\mu$g/mL, Calbiochem, Merck Millipore, Germany) and anti-KRAS (IgG rabbit polyclonal antibody, diluted 1:250, ab 102007, Abcam, United Kingdom). Membranes with the samples were incubated overnight at 4 °C with primary antibodies. The filters were washed with 0.05% Tween in PBS and subsequently incubated for 1 h with the secondary antibodies, horseradish peroxidase conjugated: antimouse IgM (diluted 1:2000) and antirabbit IgG (diluted 1:5000) (Calbiochem, Merck Millipore, Germany). The signal of the proteins was developed with Super Signal West PICO and FEMTO (ThermoFisher) and detected with ChemiDOC XRS, Quantity One 4.6.5 software (BioRad Laboratories).

**RNA Extraction and Quantitative Real-Time PCR.** Panc-1 cells were plated in a 96-well plate ($18 \times 10^3$ cells/well). The following day (24 h), we treated the cells with **2a**, **2b**, **1a**, and **1b**, and total RNA was extracted by using iScript RT-qPCR Sample Preparation Reagent (BioRad) 6 and 24 h after treatment, following the manufacturer's instructions. For the cDNA synthesis, 1.25 $\mu$L of RNA was heated at 70 °C and placed on ice. The solution was added to 11.5 $\mu$L of a mix containing: 1x buffer, 0.01 M DTT (Invitrogen, Carlsbad, CA), 1.6 $\mu$M primer dT [MWG Biotech, Germany; $d(T)_{16}$], 1.6 $\mu$M random hexamer primers (Mycrosynth, Switzerland), 0.4 mM dNTPs solution containing equimolar amounts of dATP, dCTP, dGTP, and dTTP (Euroclone, Italy), 0.8 U/$\mu$L RNase OUT, and 8 U/$\mu$L of M-MLV reverse transcriptase (Invitrogen). The mixtures were incubated for 1 h at 37 °C and stopped by heating at 95 °C for 5 min. As a negative control, the reverse transcription reaction was performed with a sample containing DEPC-treated water.

To determine the levels of *KRAS* and housekeeping genes hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) and $\beta$2-microglobulin, quantitative real-time multiplex reactions were performed. We used 1x Kapa Probe fast qPCR kit (KAPA Biosystems, Wilmington, MA), 2.2 $\mu$L of cDNA, and primers/probes (sequences are reported in Supporting Information S8). The PCR cycle was 3 min at 95 °C, 50 cycles 10 s at 95 °C, and 60 s at 58 °C. PCR reactions were carried out with a CFX-96 real-time PCR apparatus controlled by an Optical System software (version 3.1) (Bio-Rad Laboratories). The *KRAS* transcript level was normalized with the housekeeping genes.



DOI: 10.1021/acs.jmedchem.7b00622
J. Med. Chem. 2017, 60, 9448−9461

**Dual Luciferase Assays.** Panc-1 cell were plated ($20 \times 10^3$) and the following day treated with increasing concentrations of compounds **2a**, **2b**, **1a**, and **1b** (0.3−1.6 $\mu$M). After 6 h of treatment, the cells were transfected with the plasmids. The 192 nt of *KRAS* 5′UTR (NM_033360) was cloned between Hind III and Nhe I of pRL-CMV plasmid (GenScript, Piscataway, NJ). Transfection was performed by mixing vectors (10 ng/well) p-light-SwitchGear KRAS (SwitchGear Genomics, Carlsbad, CA) (in the text pRL-KRAS) or pRL-CMV-UTR or pRL-CMV (*Renilla* luciferase) with 200 ng of control plasmid (pHRAS-mutA-luc) (*Firefly* luciferase) using jet-PEI (Polyplus, France) as transfectant agent. *Renilla* luciferase in cell lysates was measured and normalized by *Firefly* luciferase. Luciferase assays were performed 48 h after transfection with Dual-Glo Luciferase Assay System (Promega, Fitchburg, WI) following the supplier instructions. Samples were read on a Turner Luminometer and the relative luminescence expressed as (T/C × 100) where T = *Renilla* luciferase/*Firefly* luciferase in treated cells and C = *Renilla* luciferase/*Firefly* luciferase in untreated cells.

**Uptake Analysis.** Panc-1 cells were plated in a 24-well plate at density of $5 \times 10^4$ cells/well. After 1 day, the cells were treated with the molecules: time and concentration are indicated in figure captions. The cells were trypsinized and pelleted. The pellets were resuspended in 200 $\mu$L of PBS and immediately analyzed by FACScan Flow Cytometer (Becton Dickinson, Franklin Lakes, NJ) equipped with a 488 nm argon laser. A minimum of $10^4$ cells for each sample were acquired in list mode and analyzed using Cell Quest software. The cell population was analyzed by FSC light and SSC light. The signal was detected by FL3 (680 nm) channel in log scale.

**Apoptosis Assays.** Caspase activity assay was performed with Apo-ONE Homogeneous Caspase-3/7 Assay (Promega,), according to the manufacturer's protocol. Annexin V−propidium iodide assay was performed with Annexin V Apoptosis Detection Kit (Santa Cruz Biotechnology, Inc., Dallas, TX), following the manufacturer's instructions. Flow cytometry measurements were performed with FACScan Flow Cytomer (Becton Dickinson).

**Biotin−Streptavidin Pull-Down Experiments.** To construct the calibration plot for biotin−streptavidin pull-down assays, we prepared several mixtures of 115- and 89-nt sequences in 50 mM Tris-HCl, pH 7.4, 100 mM KCl, at a total concentration of 0.2 $\mu$M and 115-nt/89-nt ratio varying from 1 to 5000. The mixtures have been amplified with KAPA2G Robust HotStart PCR Kit (KAPA Biosystems) using the same couple of primers (sequences reported in Supporting Information S8) (0.5 $\mu$M) and dNTPs (0.2 mM). The program used was 5 min at 95 °C, 15 s at 95 °C, 15 s at 55 °C, and 30 s at 72 °C, for 35 cycles. The products have been separated by 8% PAGE and the bands stained with ethidium bromide. Their intensities were measured with a ChemiDOC XRS, Quantity One 4.6.5 software (BioRad Laboratories). We reported in a plot the ratio of the intensities of the amplified bands (115-nt/89-nt) as a function of the logarithm of the ratio ($r$) of the two sequences in the mixtures. We obtained a straight line that correlated the band intensities with the mixture composition.

Total RNA was extracted from Panc-1 cells, and its concentration was measured by UV absorption. Eight micrograms of cellular RNA in 50 mM KCl, 50 mM Tris-HCl, pH 7.4, was incubated overnight with increasing concentrations of b-**2a** (from 0.4 to 0.8 $\mu$M). We incubated the magnetic beads, after saturation with ssDNA (vide infra), with cellular RNA treated with the biotinylated ligand for 20 min at 25 °C. The supernatant was removed, and the beads were washed twice with Tris buffer. We then recovered the bound RNA with a solution 0.8 M NaCl. The recovered RNA was retrotranscribed with 0.8 U/$\mu$L RNase OUT and 8 U/$\mu$L of M-MLV reverse transcriptase (Life Technologies, ThermoFisher) and amplified by quantitative real-time PCR (vide infra). We amplified *KRAS* and the housekeeping genes *HPRT* and $\beta$2-microglobulin (for primers, see Supporting Information S8).

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.7b00622.

Scheme illustrating how RG4 can regulate translation initiation, thermal difference spectra of RG4s, *KRAS* 5′-UTR sequences in mammals, fluorescence titrations, UV and NMR spectra and HPLC-chromatograms, confocal microscopy, colony forming assay, and DNA and RNA sequences used in the study (PDF)

SMILES data (CSV)

## ■ AUTHOR INFORMATION


### Corresponding Author

*Tel: +39.0432.494395. Fax: +39.0432.494301. E-mail: luigi.xodo@uniud.it.

### ORCID

Giovanni Capranico: 0000-0002-8708-6454

Luigi E. Xodo: 0000-0003-3344-7207


### Author Contributions

This study was conceived and written by L.E.X. and G.M. A.S.T. and A.S. designed and synthesized the G4 ligands, and G.M. and S.C. performed the experiments. J.M. and G.C. prepared and purified BG4. All authors have given approval to the final version of the manuscript.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work has been carried out with the financial support of AIRC (Italian association for cancer research). IG2013, project code 143.

## ■ ABBREVIATIONS USED

*KRAS*, Kirsten ras gene; PDAC, pancreatic ductal adenocarcinoma cells; RG4, RNA G-quadruplex; Panc-1, pancreatic cancer cells; BxPC-3, pancreatic cancer cells; ATPD, antrathiophenediones; 5′-UTR, 5′-untranslated region; ATFD, antrafurandiones; DCM, dichloromethane; HATU, 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]-pyridinium 3-oxide hexafluorophosphate; NMM, *N*-methylmorpholine; PyBOP, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate; DIPEA, ethyldiisopropylamine; THF, tetrahydrofuran; TFA, trifluoroacetic acid; ThT, thioflavin T

## ■ REFERENCES


(1) Vigil, D.; Cherfils, J.; Rossman, K. L.; Der, C. J. Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? *Nat. Rev. Cancer* **2010**, *10*, 842−857.

(2) Rodriguez-Viciana, P.; Sabatier, C.; McCormick, F. Signaling specificity by Ras family GTPases is determined by the full spectrum of effectors they regulate. *Mol. Cell. Biol.* **2004**, *24*, 4943−4945.

(3) Wortzel, I.; Seger, R. The ERK cascade: distinct functions within various subcellular organelles. *Genes Cancer* **2011**, *2*, 195−209.

(4) Castellano, E.; Downward, J. Ras interaction with PI3K: more than just another effector pathway. *Genes Cancer* **2011**, *2*, 261−274.

(5) Schubbert, S.; Shannon, K.; Bollag, G. Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer* **2007**, *7*, 295−308.




DOI: 10.1021/acs.jmedchem.7b00622
*J. Med. Chem.* 2017, 60, 9448−9461

(6) Ryan, D. P.; Hong, T. S.; Bardeesy, N. Pancreatic adenocarcinoma. *N. Engl. J. Med.* **2014**, *371*, 2140−2141.
(7) Prior, I. A.; Lewis, P. D.; Mattos, C. A comprehensive survey of Ras mutations in cancer. *Cancer Res.* **2012**, *72*, 2457−2467.
(8) Karnoub, A. E.; Weinberg, R. A. Ras oncogenes: split personalities. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 517−531.
(9) Collins, M. A.; Bednar, F.; Zhang, Y.; Brisset, J. C.; Galbán, S.; Galbán, C. J.; Rakshit, S.; Flannagan, K. S.; Adsay, N. V.; Pasca di Magliano, M. Oncogenic Kras is required for both the initiation and maintenance of pancreatic cancer in mice. *J. Clin. Invest.* **2012**, *122*, 639−653.
(10) Kanda, M.; Matthaei, H.; Wu, J.; Hong, S. M.; Yu, J.; Borges, M.; Hruban, R. H.; Maitra, A.; Kinzler, K.; Vogelstein, B.; Goggins, M. Presence of somatic mutations in most early-stage pancreatic intraepithelial neoplasia. *Gastroenterology* **2012**, *142*, 730−733.
(11) Vincent, A.; Herman, J.; Schulick, R.; Hruban, R. H.; Goggins, M. Pancreatic cancer. *Lancet* **2011**, *378*, 607−620.
(12) Ying, H.; Kimmelman, A. C.; Lyssiotis, C. A.; Hua, S.; Chu, G. C.; Fletcher-Sananikone, E.; Locasale, J. W.; Son, J.; Zhang, H.; Coloff, J. L.; Yan, H.; Wang, W.; Chen, S.; Viale, A.; Zheng, H.; Paik, J. H.; Lim, C.; Guimaraes, A. R.; Martin, E. S.; Chang, J.; Hezel, A. F.; Perry, S. R.; Hu, J.; Gan, B.; Xiao, Y.; Asara, J. M.; Weissleder, R.; Wang, Y. A.; Chin, L.; Cantley, L. C.; DePinho, R. A. Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* **2012**, *149*, 656−670.
(13) Son, J.; Lyssiotis, C. A.; Ying, H.; Wang, X.; Hua, S.; Ligorio, M.; Perera, R. M.; Ferrone, C. R.; Mullarky, E.; Shyh-Chang, N.; Kang, Y.; Fleming, J. B.; Bardeesy, N.; Asara, J. M.; Haigis, M. C.; DePinho, R. A.; Cantley, L. C.; Kimmelman, A. C. Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* **2013**, *496*, 101−105.
(14) Weinstein, I. B.; Joe, A.; Felsher, D. Oncogene addiction. *Cancer Res.* **2008**, *68*, 3077−3080.
(15) Patricelli, M. P.; Janes, M. R.; Li, L. S.; Hansen, R.; Peters, U.; Kessler, L. V.; Chen, Y.; Kucharski, J. M.; Feng, J.; Ely, T.; Chen, J. H.; Firdaus, S. J.; Babbar, A.; Ren, P.; Liu, Y. Selective inhibition of oncogenic KRAS output with small molecules targeting the inactive state. *Cancer Discovery* **2016**, *6*, 316−329.
(16) Wilson, C. Y.; Tolias, P. Recent advances in cancer drug discovery targeting RAS. *Drug Discovery Today* **2016**, *21*, 1915−1919.
(17) Stephen, A. G.; Esposito, D.; Bagni, R. K.; McCormick, F. Dragging ras back in the ring. *Cancer Cell* **2014**, *25*, 272−281.
(18) Ostrem, J. M.; Shokat, K. M. Direct small-molecule inhibitors of KRAS: from structural insights to mechanism-based design. *Nat. Rev. Drug Discovery* **2016**, *15*, 771−785.
(19) Kiessling, M. K.; Rogler, G. Targeting the ras pathwayby mitogene-activated protein kinase inhibitors. *Swiss Med. Wkly.* **2015**, *145*, w14207.
(20) Li, T.; Sparano, J. A. Farnesyl transferase inhibitors. *Cancer Invest.* **2008**, *26*, 653−661.
(21) Cox, A. D.; Der, C. J.; Philips, M. R. Targeting RAs membrane association: back to the future for anti-RAS drug discovery. *Clin. Cancer Res.* **2015**, *21*, 1819−1827.
(22) Tamanoi, F.; Lu, J. Recent progress in developing small molecule inhibitors designed to interfere with ras membrane association: toward inhibiting KRAS and NRAS functions. *Enzymes* **2013**, *34*, 181−200.
(23) Cogoi, S.; Paramasivam, M.; Spolaore, B.; Xodo, L. E. Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins. *Nucleic Acids Res.* **2008**, *36*, 3765−3780.
(24) Faudale, M.; Cogoi, S.; Xodo, L. E. Photoactivated cationic alkyl-substituted porphyrin binding to g4-RNA in the 5′-UTR of KRAS oncogene represses translation. *Chem. Commun. (Cambridge, U. K.)* **2012**, *48*, 874−876.
(25) Rapozzi, V.; Zorzet, S.; Zacchigna, M.; Della Pietra, E.; Cogoi, S.; Xodo, L. E. Anticancer activity of cationic porphyrins in melanoma tumour-bearing mice and mechanistic in vitro studies. *Mol. Cancer* **2014**, *13*, 75−92.
(26) Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. An RNA G-quadruplex in the 5′-UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.* **2007**, *3*, 218−221.
(27) Bugaut, A.; Balasubramanian, S. 5′-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res.* **2012**, *40*, 4727−4741.
(28) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593−11598.
(29) Bochman, M. L.; Paeschke, K.; Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **2012**, *13*, 770−780.
(30) Murat, P.; Balasubramanian, S. Existance and consequences of G-quadruplex structures in DNA. *Curr. Opin. Genet. Dev.* **2014**, *25*, 22−29.
(31) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **2006**, *34*, 5402−5415.
(32) Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discovery* **2011**, *10*, 261−275.
(33) Maizels, N. G4 motifs in human genes. *Ann. N. Y. Acad. Sci.* **2012**, *1267*, 53−60.
(34) Mendoza, O.; Bourdoncle, A.; Boulé, J. B.; Brosh, R. M., Jr.; Mergny, J. L. G-quadruplexes and helicases. *Nucleic Acids Res.* **2016**, *44*, 1989−2006.
(35) Cogoi, S.; Xodo, L. E. G4 DNA in ras genes and its potential in cancer therapy. *Biochim. Biophys. Acta, Gene Regul. Mech.* **2016**, *1859*, 663−674.
(36) Hänsel-Hertsch, R.; Beraldi, D.; Lensing, S. V.; Marsico, G.; Zyner, K.; Parry, A.; Di Antonio, M.; Pike, J.; Kimura, H.; Narita, M.; Tannahill, D.; Balasubramanian, S. G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.* **2016**, *48*, 1267−1272.
(37) Beaudoin, J. D.; Perreault, J. P. 5′-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res.* **2010**, *38*, 7022−7036.
(38) Jodoin, R.; Bauer, L.; Garant, J. M.; Mahdi Laaref, A.; Phaneuf, F.; Perreault, J. P. The folding of 5′-UTR human G-quadruplexes possessing long central loop. *RNA* **2014**, *20*, 1129−1141.
(39) Hinnebusch, A. G.; Ivanov, I. P.; Sonenberg, N. Translational control by 5′-untranslated regions of eukaryotic mRNAs. *Science* **2016**, *352*, 1413−1416.
(40) Song, J.; Perreault, J. P.; Topisirovic, I.; Richard, S. RNA G-quadruplexes and their potential regulatory roles in translation. *Translation* **2016**, *4*, No. e1244031.
(41) Gomez, D.; Guédin, A.; Mergny, J. L.; Salles, B.; Riou, J. F.; Teulade-Fichou, M. P.; Calsou, P. A G-quadruplex structure within the 5′-UTR of TRF2 mRNA represses translation in human cells. *Nucleic Acids Res.* **2010**, *38*, 7187−7198.
(42) Kwok, C. K.; Ding, Y.; Shahid, S.; Assmann, S. M.; Bevilacqua, P. C. A stable RNA G-quadruplex within the 5′-UTR of Arabidopsis thaliana ATR mRNA inhibits translation. *Biochem. J.* **2015**, *467*, 91−102.
(43) Arora, A.; Dutkiewicz, M.; Scaria, V.; Hariharan, M.; Maiti, S.; Kurreck, J. Inhibition of translation in eukaryotic cells by an RNA G-quadruplex motif. *RNA* **2008**, *14*, 1290−1296.
(44) Kumari, S.; Bugaut, A.; Balasubramanian, S. Position and stability are determining factors for translation repression by an RNA G-quaduplex forming sequence within 5′-UTR of the NRAS proto oncogene. *Biochemistry* **2008**, *47*, 12664−12669.
(45) Morris, M. J.; Negishi, Y.; Pazsint, C.; Schonhoft, J. D.; Basu, S. A RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES. *J. Am. Chem. Soc.* **2010**, *132*, 17831−17839.
(46) Lammich, S.; Kamp, F.; Wagner, J.; Nuscher, B.; Zilow, S.; Ludwig, A. K.; Willem, M.; Haass, C. Translational repression of the disintegrin and metallo protease ADAM10 by a stable G-quadruplex


DOI: 10.1021/acs.jmedchem.7b00622
*J. Med. Chem.* 2017, 60, 9448−9461

secondary structure in its 5′-untrasnlated region. *J. Biol. Chem.* **2011**, *286*, 45063−45072.
(47) Bolduc, F.; Garant, J. M.; Allard, F.; Perreault, J. P. Irregular G-quadruplexes found in untranslated regions of human mRNAs influences translation. *J. Biol. Chem.* **2016**, *291*, 21751−21750.
(48) Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003**, *31*, 3406−3415.
(49) Kikin, O.; D'Antonio, L.; Bagga, P. S. QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* **2006**, *34*, W676−W682.
(50) Vorlickova, M.; Kejnovska, I.; Bednarova, K.; Renciuk, D.; Kypr, J. Circular dichroism spectroscopy of DNA: from duplexes to quadruplexes. *Chirality* **2012**, *24*, 691−698.
(51) Mergny, J. L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.* **2005**, *33*, No. e138.
(52) Yangyuoru, P. M.; Di Antonio, M.; Ghimire, C.; Biffi, G.; Balasubramanian, S.; Mao, H. Dual binding of an antibody and a small molecule increases the stability of TERRA G-quadruplex. *Angew. Chem., Int. Ed.* **2015**, *54*, 910−913.
(53) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **2013**, *5*, 182−186.
(54) Miserachs, H. G.; Donghi, D.; Borner, R.; Johannsen, S.; Sigel, R. K. O. Distinct differences in metal ion specificity of RNA and DNA G-quadruplexes. *JBIC, J. Biol. Inorg. Chem.* **2016**, *21*, 975−986.
(55) McPike, M. P.; Goodisman, J.; Dabrowiak, J. C. Drug-RNA footprinting. *Methods Enzymol.* **2001**, *340*, 431−449.
(56) Xu, S.; Li, Q.; Xiang, J.; Yang, Q.; Sun, H.; Guan, A.; Wang, L.; Liu, Y.; Yu, L.; Shi, Y.; Chen, H.; Tang, Y. Thioflavin T as an efficient fluorescence sensor for selective recognition of RNA G-quadruplexes. *Sci. Rep.* **2016**, *6*, 24793.
(57) Weldon, C.; Behm-Ansmant, I.; Hurley, L. H.; Burley, G. A.; Branlant, C.; Eperon, I. C.; Dominguez, C. Identification of G-quadruplexes in long functional RNAs using 7-deazaguanine RNA. *Nat. Chem. Biol.* **2017**, *13*, 18−20.
(58) Bugaut, A.; Rodriguez, R.; Kumari, S.; Hsu, S. T.; Balasubramanian, S. Small molecule-mediated inhibition of translation by targeting a native RNA G-quadruplex. *Org. Biomol. Chem.* **2010**, *8*, 2771−2776.
(59) Katsuda, Y.; Sato, S.; Asano, L.; Morimura, Y.; Furuta, T.; Sugiyama, H.; Hagihara, M.; Uesugi, M. A small molecule that represses translation of G-quadruplex-containing mRNA. *J. Am. Chem. Soc.* **2016**, *138*, 9037−9040.
(60) Cogoi, S.; Shchekotikhin, A. E.; Membrino, A.; Sinkevich, Y. B.; Xodo, L. E. Guanidino anthrathiophenediones as G-quadruplex binders: uptake, intracellular localization, and anti-Harvey-Ras gene activity in bladder cancer cells. *J. Med. Chem.* **2013**, *56*, 2764−2778.
(61) Cogoi, S.; Zorzet, S.; Shchekotikhin, A. E.; Xodo, L. E. Potent apoptotic response induced by chloroacetamidine anthrathiophenediones in bladder cancer cells. *J. Med. Chem.* **2015**, *58*, 5476−5485.
(62) Shchekotikhin, A. E.; Glazunova, V. A.; Dezhenkova, L. G.; Luzikov, Y. N.; Sinkevich, Y. B.; Kovalenko, L. V.; Buyanov, V. N.; Balzarini, J.; Shtil, A. A.; Preobrazhenskaya, M. N.; et al. Synthesis and cytotoxic properties of 4,11-bis[(aminoethyl)amino]anthra[2,3-*b*]-thiophene-5,10-diones, novel analogues of antitumor anthracene-9,10-diones. *Bioorg. Med. Chem.* **2009**, *17*, 1861−1869.
(63) Tikhomirov, A. S.; Shchekotikhin, A. E.; Preobrazhenskaya, M. N. Methods for the synthesis and modification of linear anthrafurandiones. *Chem. Heterocycl. Compd.* **2014**, *50*, 171−184.
(64) Gronowitz, S. Thiophene and Its Derivatives, Part 2. *The Chemistry of Heterocyclic Compounds*; John Wiley & Sons: New York, 1986; Vol. *44*.
(65) Tikhomirov, A. S.; Shchekotikhin, A. E.; Lee, Y. H.; Chen, Y. A.; Yeh, C. A.; Tatarskiy, V. V.; Dezhenkova, L. G.; Glazunova, V. A.; Balzarini, J.; Shtil, A. A.; Preobrazhenskaya, M. N.; Chueh, P. J. Synthesis and characterization of 4,11-diaminoanthra[2,3-*b*]furan-5,10-diones: tumor cell apoptosis through tNOX-modulated $NAD^+$/NADH ratio and SIRT1. *J. Med. Chem.* **2015**, *58*, 9522−9534.
(66) Sabatino, G.; Chinol, M.; Paganelli, G.; Papi, S.; Chelli, M.; Leone, G.; Papini, A. M.; De Luca, A.; Ginanneschi, M. A new biotin derivative-DOTA conjugate as a candidate for pretargeted diagnosis and therapy of tumors. *J. Med. Chem.* **2003**, *46*, 3170−3173.
(67) Huang, F.; He, J.; Zhang, Y.; Guo, Y. Synthesis of biotin−AMP conjugate for 5′ biotin labeling of RNA through one-step in vitro transcription. *Nat. Protoc.* **2008**, *3*, 1848−1861.
(68) Isaac-Lam, M. F.; Hammonds, D. M. Biotinylated chlorin and its zinc and indium complexes: synthesis and in vitro biological evaluation for photodynamic therapy. *Pharmaceuticals* **2017**, *10*, 41.
(69) Tikhomirov, A. S.; Shchekotikhin, A. E.; Luzikov, Y. N.; Korolev, A. M.; Preobrazhenskaya, M. N. Heterocyclic analogs of 5,12-naphthacene-Quinone. 11. A new method for preparing 4,11-dimethoxyanthra[2,3-*b*]furan-5,10-dione. *Chem. Heterocycl. Compd.* **2011**, *47*, 1206−1211.
(70) Phan, A. T. Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J.* **2010**, *277*, 1107−1117.
(71) Cogoi, S.; Shchekotikhin, A. E.; Xodo, L. E. HRAS is silenced by two neighboring G-quadruplexes and activated by MAZ, a zinc-finger transcription factor with DNA unfolding property. *Nucleic Acids Res.* **2014**, *42*, 8379−8388.
(72) Cogoi, S.; Rapozzi, V.; Cauci, S.; Xodo, L. E. Critical role of hnRNP A1 in activating KRAS transcription in pancreatic cancer cells: A molecular mechanism involving G4 DNA. *Biochim. Biophys. Acta, Gen. Subj.* **2017**, *1861*, 1389−1398.
(73) Cogoi, S.; Xodo, L. E. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res.* **2006**, *34*, 2536−2549.
(74) Cogoi, S.; Zorzet, S.; Rapozzi, V.; Géci, I.; Pedersen, E. B.; Xodo, L. E. MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic Acids Res.* **2013**, *41*, 4049−4064.
(75) Wlodkowic, D.; Skommer, J.; Darzynkiewicz, Z. Flow cytometry-based apoptosis detection. *Methods Mol. Biol.* **2009**, *559*, 19−32.
(76) Biffi, G.; Di Antonio, M.; Tannahill, D.; Balasubramanian, S. Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat. Chem.* **2014**, *6*, 75−80.
(77) Guo, J. U.; Bartel, D. P. RNA G-quadruplexes are globally unfolded in eukaryotiv cells and depleted in bacteria. *Science* **2016**, *353*, 6306.
(78) Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expression Purif.* **2005**, *41*, 207−234.



DOI: 10.1021/acs.jmedchem.7b00622
*J. Med. Chem.* 2017, 60, 9448−9461

# Supporting Information

# RNA G-quadruplexes in Kirstein ras (*KRAS*) oncogene as targets for small molecules inhibiting translation

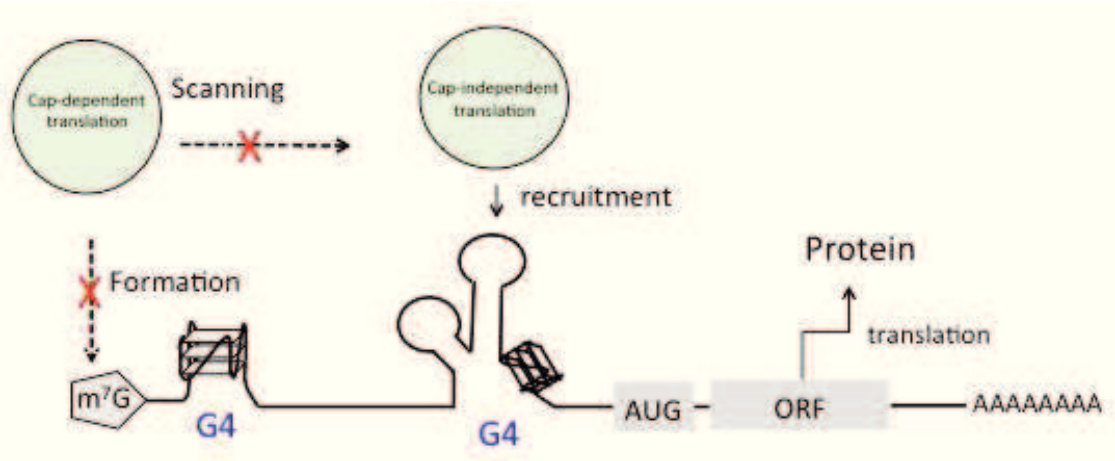
*Giulia Miglietta,* [1] *Susanna Cogoi,* [1] *Jessica Marinello,* [2] *Giovanni Capranico,* [2] *Alexander S. Tikhomirov,* [3] *Andrey Shchekotikhin,* [3] *Luigi E. Xodo* [1] ***

[1] Department of Medical and Biological Sciences, University of Udine, 33100 Udine, Italy;

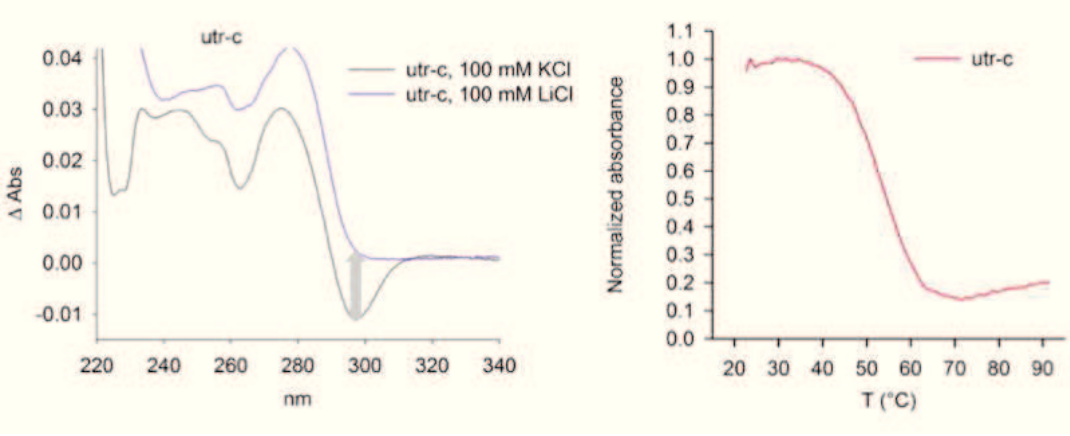
[2] Department of Pharmacy and Biotechnology, University of Bologna, 40100 Bologna, Italy;

[3] Gause Institute of New Antibiotics, 119021 Moscow, Russia;

Contents:

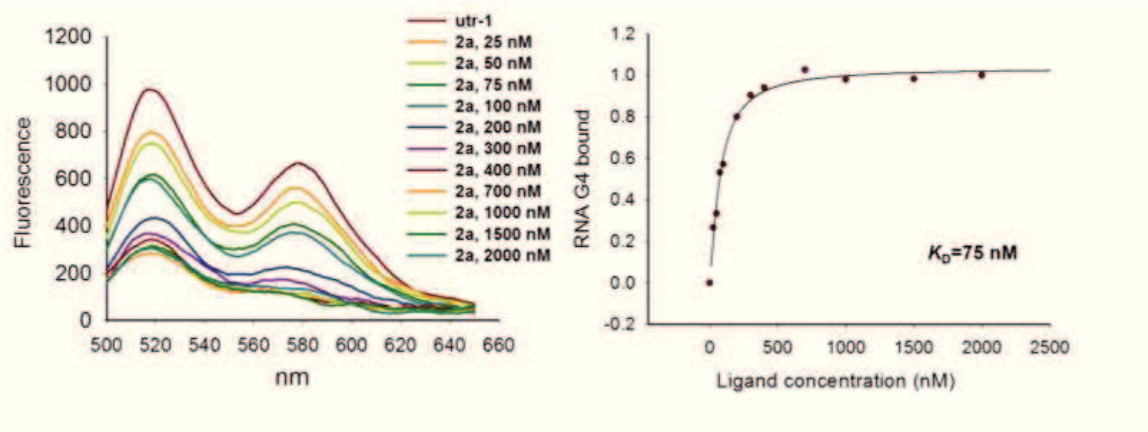



**Suppl. data S1:** Possible role played by RG4s in the regulation of translation initiation. RG4s may repress translation by inhibiting the assembly of the translation initiation complex at the 5' $m^7G$ cap of the transcript or disturbing the scanning of the initiation complex towards the AUG start codon.

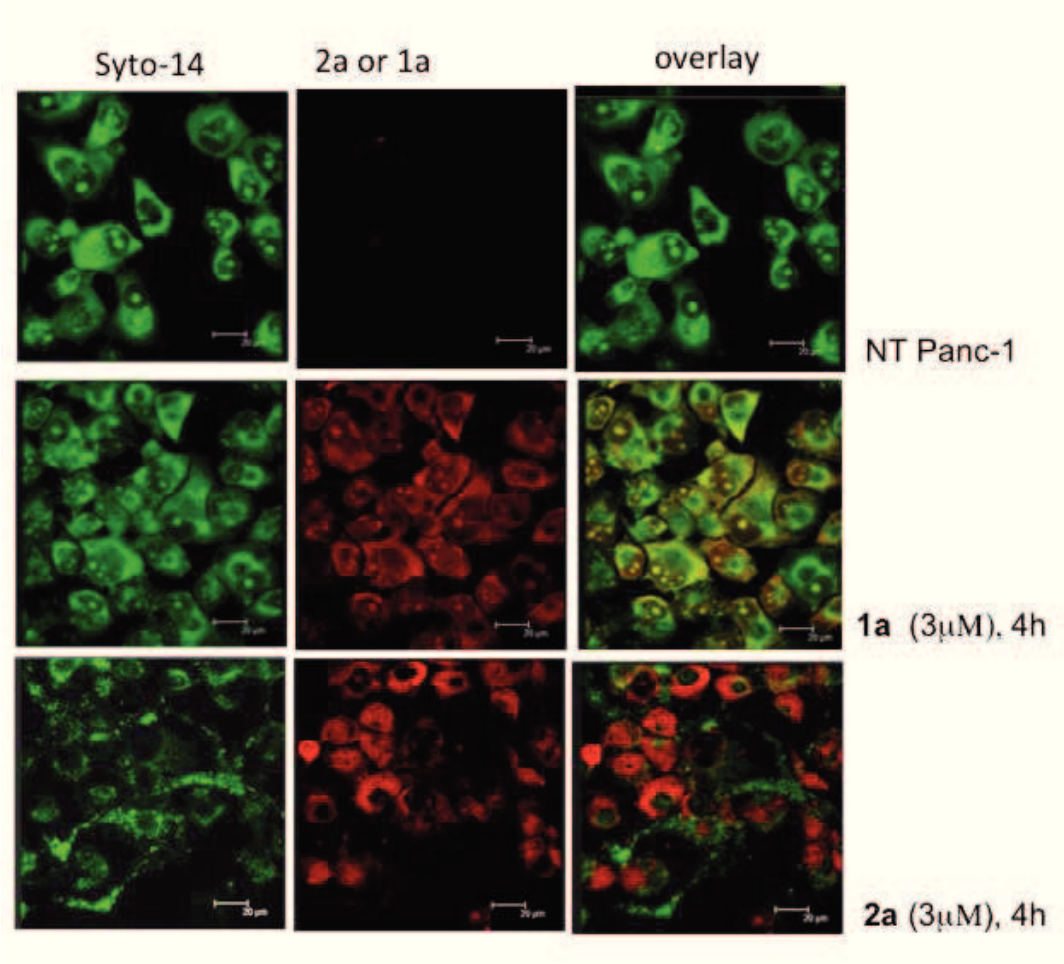



**Suppl. data S2:** Typical thermal difference spectra (TDS) (Spectrum at 90 °C – spectrum at 20 °C) for *utr-c* in $K^+$-buffer (50 mM cacodylate, pH 7.4, 100 mM KCl) and $Li^+$-buffer (50 mM cacodylate, pH 7.4, 100 mM LiCl). Right panels show the melting curve measured at 296 nm. Sequences utr-z and utr-1 show similar TDS and melting curves.

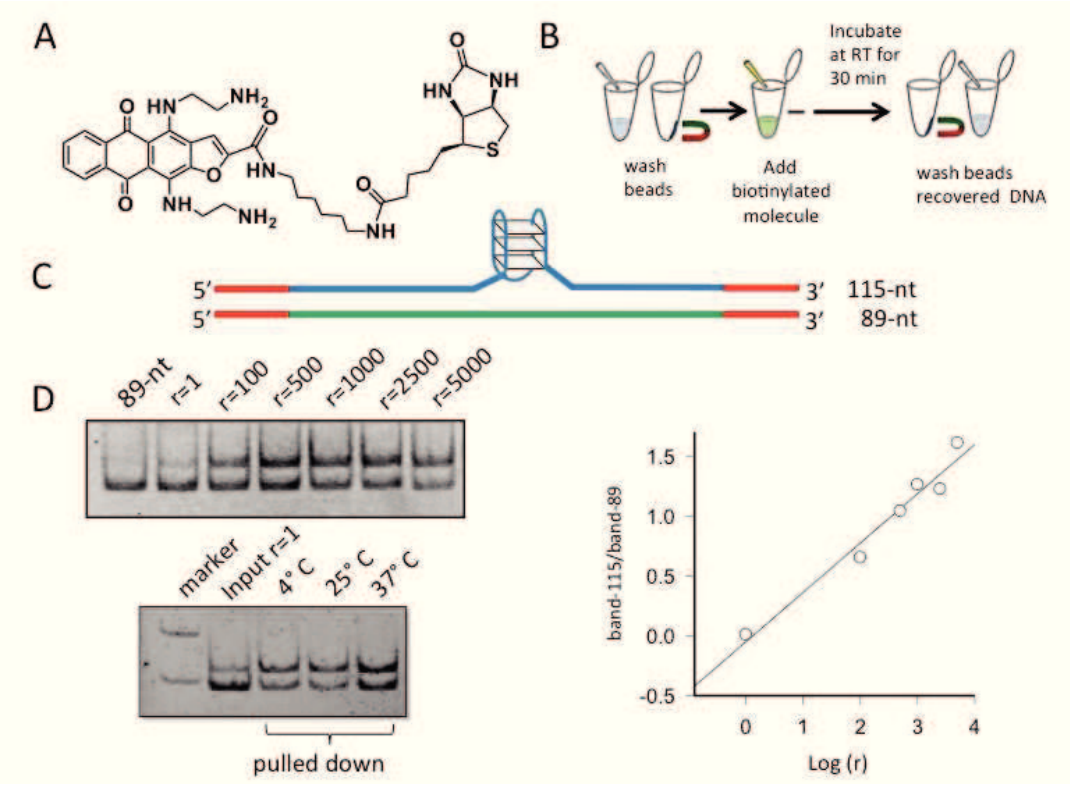

**Suppl. data, S3:** *KRAS* 5'-UTR sequences of various mammals. The 5'-UTR sequences have a high homology. The G4 motifs are indicated with coloured lines: red line (utr-z); blue line (utr-1) and green line (utr-c).




**Suppl. data, S4.** Sequence *utr*-1 was labelled at the 5'- and 3'-ends with fluorescein and rhodamine, respectively. The sequence excited at 480 nm emitted a fluorescence spectrum with two emission peaks at 520 and 580 nm. As the compound **2a** absorbs in the range between 490 and 610 nm, upon binding to utr-1 RG4, it quenches the emission of the fluorophores. From the titration spectra we obtained the binding curve that was best-fitted to a standard Hill equation. The analysis gave a $K_D$ of 75 nM, with an uncertainty of about 10 %.

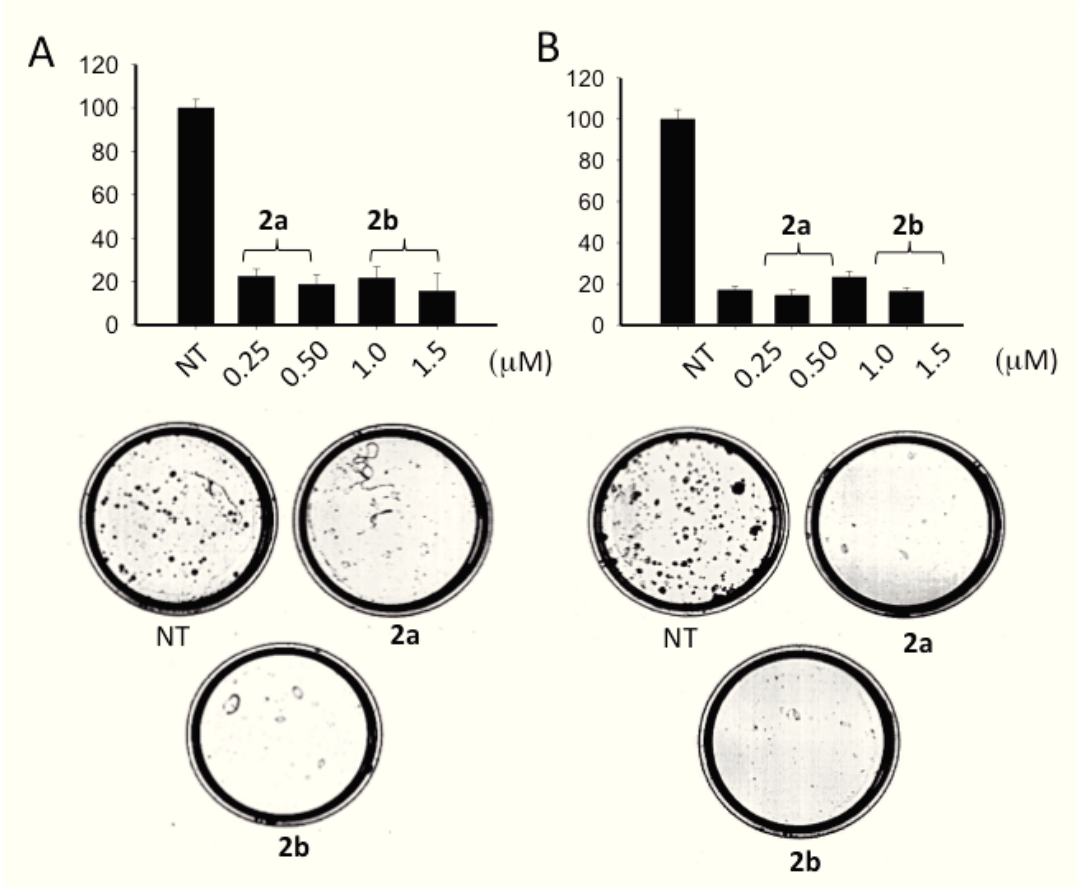



**Suppl. Data S5.** Confocal microscopy images of Panc-1 cells treated with 3 μM **2a** or **1a** for 4 h. The cells have been treated with Syto-14, a fluorescent dye that binds to nucleic acids (it has more affinity for RNA than DNA) and stains the cells in green. The G4 ligands **2a** and **1a** emit red fluorescence. The images show that these molecules are efficiently taken up by then cells and localize mainly in the cytoplasm than in the nucleus.




**Suppl. data S6.** We designed two DNA strands of 115 and 89 nt of which the former contained in the middle a G4 motif (telomeric *htel*). Both strands having the same 5'- and 3'-ends could be amplified with the same couple of primers. We amplified several mixtures of the two DNA strands (total strand concentration, [115-nt]+[89-nt]= 0.2 μM) at ratio *r* varying from 1 to 5000 (*r*=[115-nt]/[89-nt]). The amplified sequences were separated by 8 % PAGE and the intensities of the 115-nt and 89-nt bands were measured with a densitometer. Plotting the ratio of the amplified bands (band-115/band-89) as a function of log (*r*), we obtained a calibration curve that correlates *r* (the mixture composition) with the ratio of the amplified bands (band-115/band-89). Note that the amplification of mixture 1:1 (*r*=1) gave a ratio band-115/band-89 of 0.01, as G4 in the 115-nt strand stalled DNA polymerase. In a parallel experiment we incubated a 1:1 mixture (0.2 μM) with 100-fold excess b-**2a** for 1.5 h at 4, 25 and 37 °C. The DNA bound to the biotinylated ligand was pulled down with streptavidin-coated magnetic beads. The unbound DNA was washed away, while the bound DNA was recovered with 0.8 M NaCl. Input (untreated

1:1 mixture) and the recovered DNA were amplified and the products separated by 8 % PAGE, under the same experimental conditions used to obtain the calibration curve. The results showed that the pulled down DNA gave a ratio 115-nt/89-nt > 1, while the input (1:1 mixture) gave a ratio << 1. On the calibration curve, the pulled-down DNA was dramatically enriched of the 115-nt sequence compared to the input 1:1 mixture (from ~ 2000-fold at 4 °C, to ~ 4000-fold at 25 °C). This experiment clearly demonstrates that b-**2a** efficiently pulled down the DNA strand carrying G4: a prerequisite essential for the next step of the experiment.



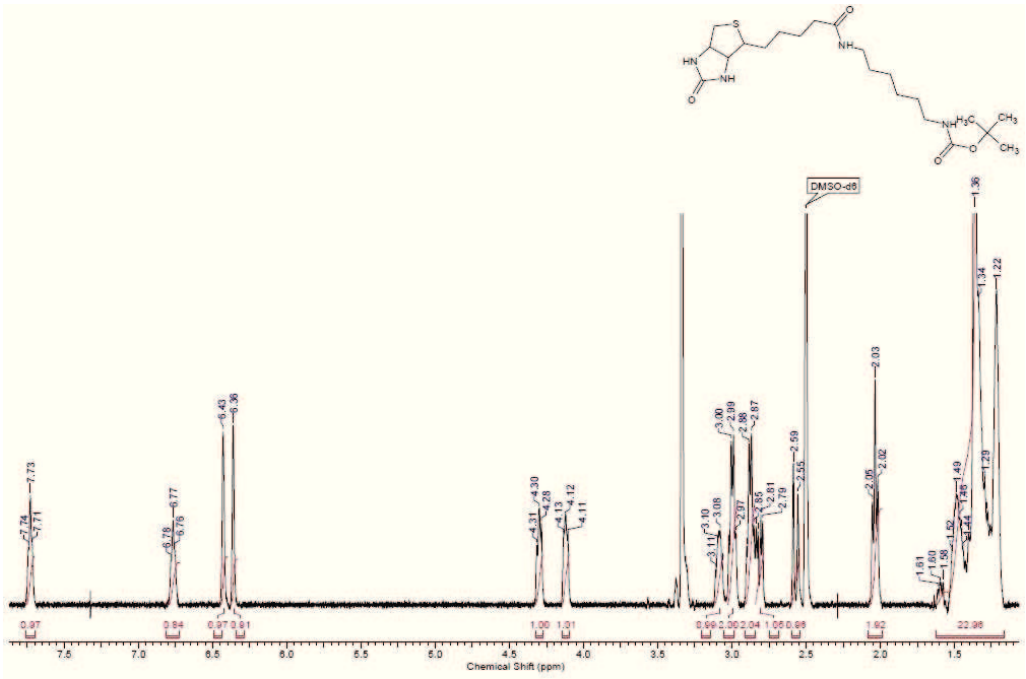

**Suppl. Data S7**. Colony-forming assay. 500 Panc-1 cells and 1000 BxPC-3 cells were plated in a 60-mm plate and treated with 0.25 and 0.5 μM **2a** or 1 and 1.5 μM **2b**. After 18 days, the cells were fixed and stained for 10 min with 2.5 % methylene blue in 50 % ethanol. Colonies with > 50 cells were counted.

**Suppl. data S8:** Sequences of primers and probes used for the qRT-PCR

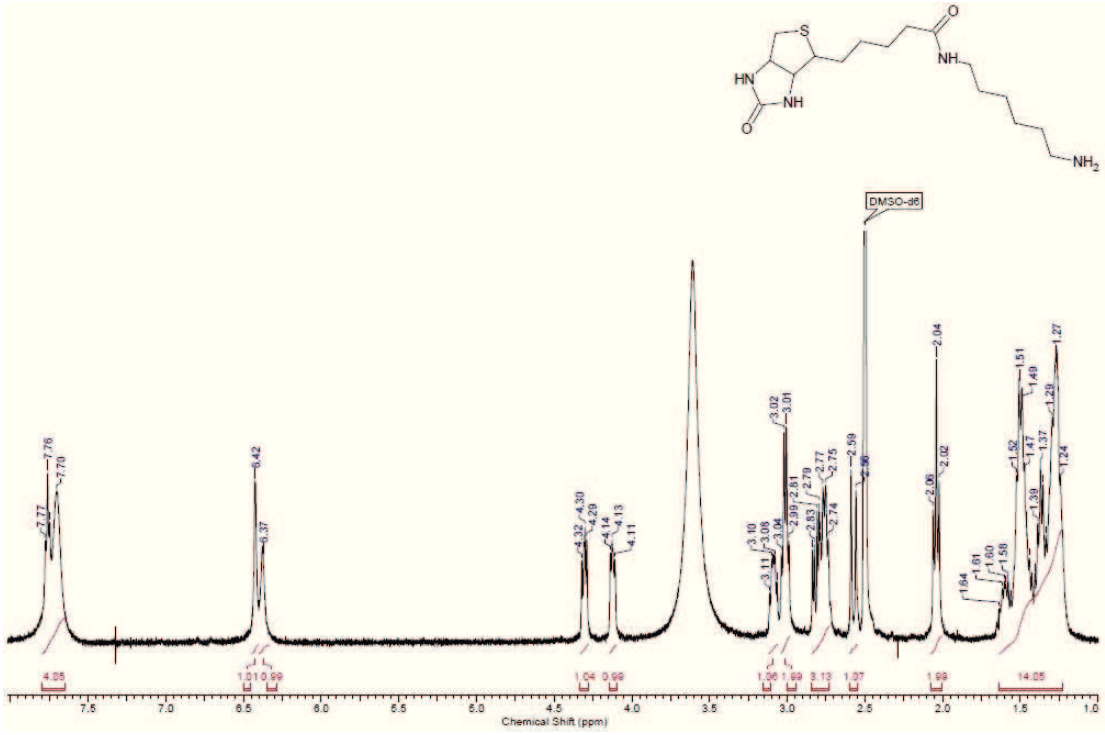
| Accession Number | SENSE<br>5' → 3' | ANTISENSE<br>5' → 3' | PROBE<br>5' → 3' |
|---|---|---|---|
| NM_033360<br>(KRAS) | CGAATATGATCCA--ACAATAGAG<br>(from 271 to 292) | ATGTACTGGTCCCTCATT<br>(from 379 to 396) | FAM-TACTCCTCTTGACCTGCTGTG-BHQ1<br>(from 352 to 372) |
| NM_004048<br>(β2-microglobulin) | CCCCACTGAAAAAGATGA<br>(from 333 to 350) | CCATGATGCTGCTTACAT<br>(from 415 to 432) | ROX-TATGCCTGCCGTGTGAACC-BHQ2<br>(from 352 to 370) |
| NM_000194<br>(HPRT) | CTTGATTGTGGAAGATATAATTG<br>(from 557 to 575) | TATATCCAACACTTCGTGG<br>(from 672 to 690) | Cy5-CTTGCGACCTTGACCATCTT-BHQ2<br>(from 633 to 652) |
| 115-nt<br>(5' → 3') | AGCTACCTTGATGAATCCAGACGGGCGTACGAGAGGGAGCGGCTGAGGGCGGTGTGGGAAGAGGGAAGAGGGGGAATCGCTACCGTTAAGCATCGATCAGATCAAGTGATAGTA | | |
| 89-nt<br>(5' → 3') | AGCTACCTTGATGAATCCAAACAAAGTCATAGGGTCAGGATGGTGCCTAATCGCTACCGTTAAGCATCGATCAGATCAAGTGATAGTAC | | |
| s-80<br>(5' → 3') | UCCUAGGCGGCGGCCGCGGCGGCGGAGGCAGCAGCGGCGGCGGCAGUGGCGGCGGCGAAGGUGGCGGCGGCUCGGCCAGU | | |
| Primer biotin pull-down sense<br>(5' → 3') | ACG TAC CTT GAT GAA TCC | | |
| Primer biotin pull-down antisense<br>5' → 3' | GTA CTA TCA CTT GAT CTG AT | | |
| RNA hairpin<br>(5' → 3') | GGCCGCCGCAGUGGCGGCGG | | |
| ssRNA<br>5' → 3' | UGUAAACAUCCUACACUCAGCU | | |

**S 8.1 $^{1}$H, $^{13}$C NMR Spectra**

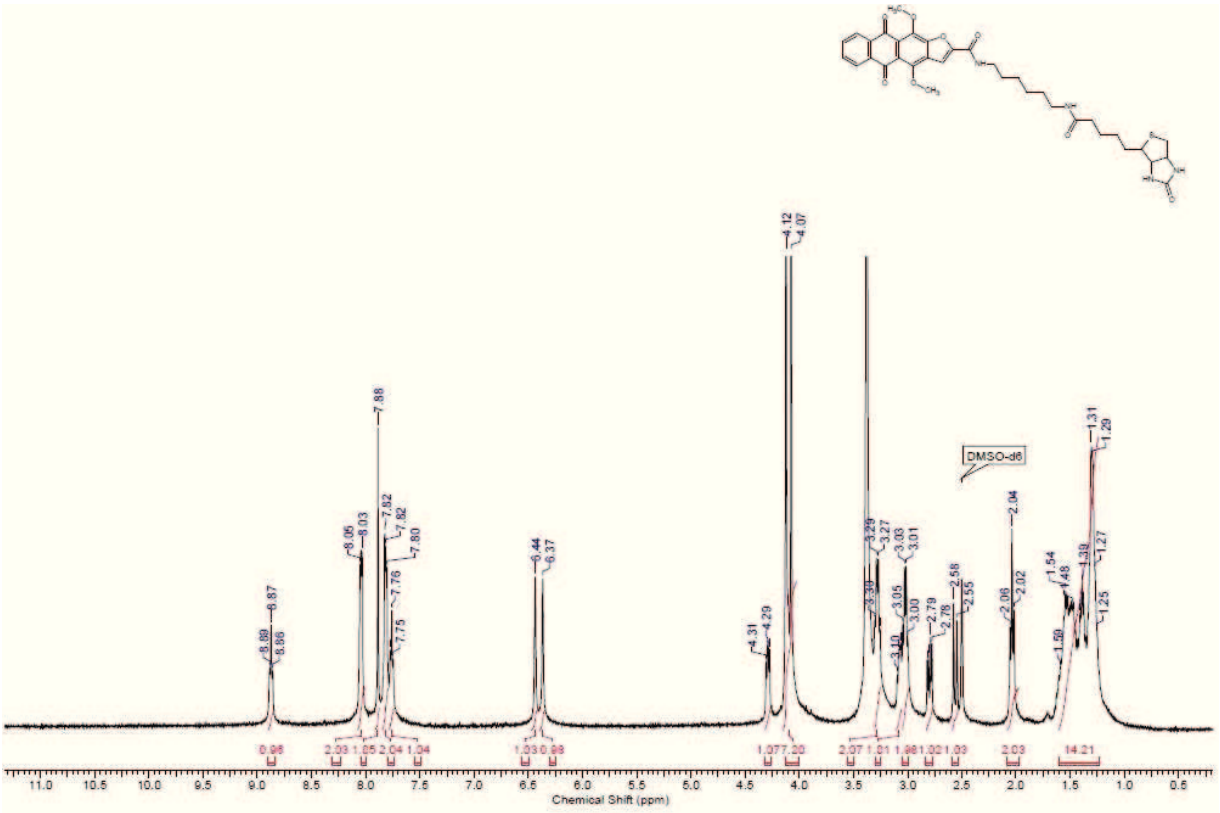
*tert-Butyl (6-(5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)-hexyl)carbamate (N-Boc-HDA-Biotin, **2**)*



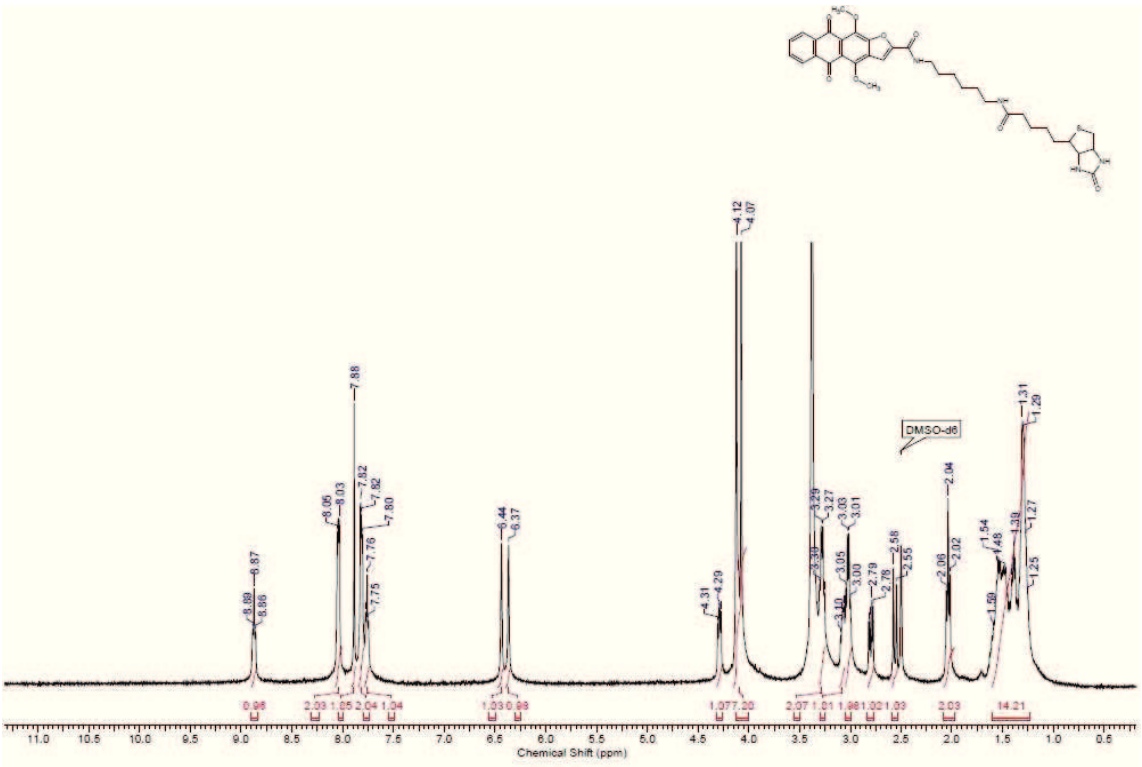
*N-(6-Aminohexyl)-5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide triflouroacetate (HDA-Biotin, **3**)*

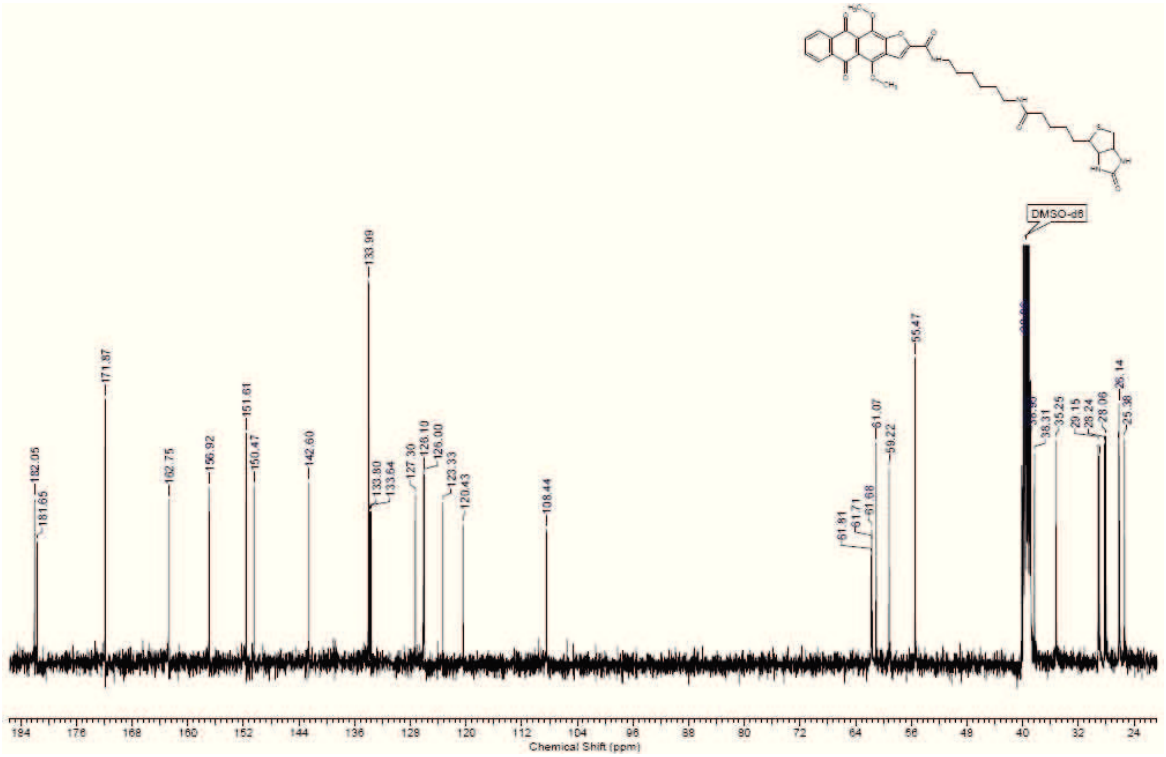


*Biotinyl-N-(6-aminohexyl)-4,11-dimethoxy-5,10-dioxo-5,10-dihydroanthra[2,3-b]furan-2-carboxamide (**4**)*

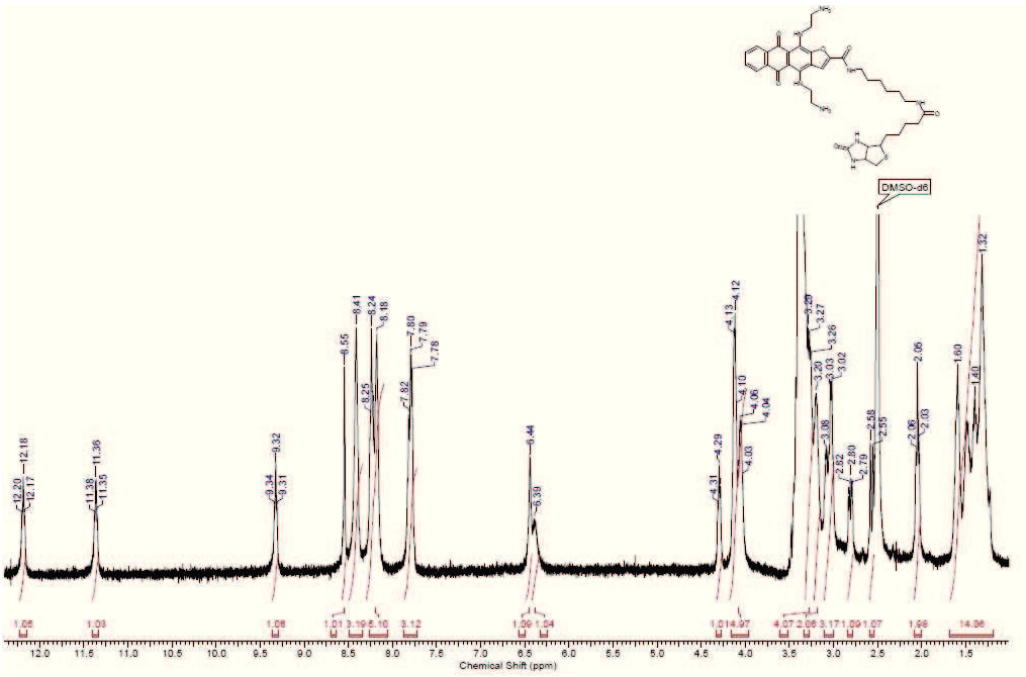


*Biotinyl-N-(6-aminohexyl)-4,11-dimethoxy-5,10-dioxo-5,10-dihydroanthra[2,3-b]furan-2-carboxamide* ***(5)***



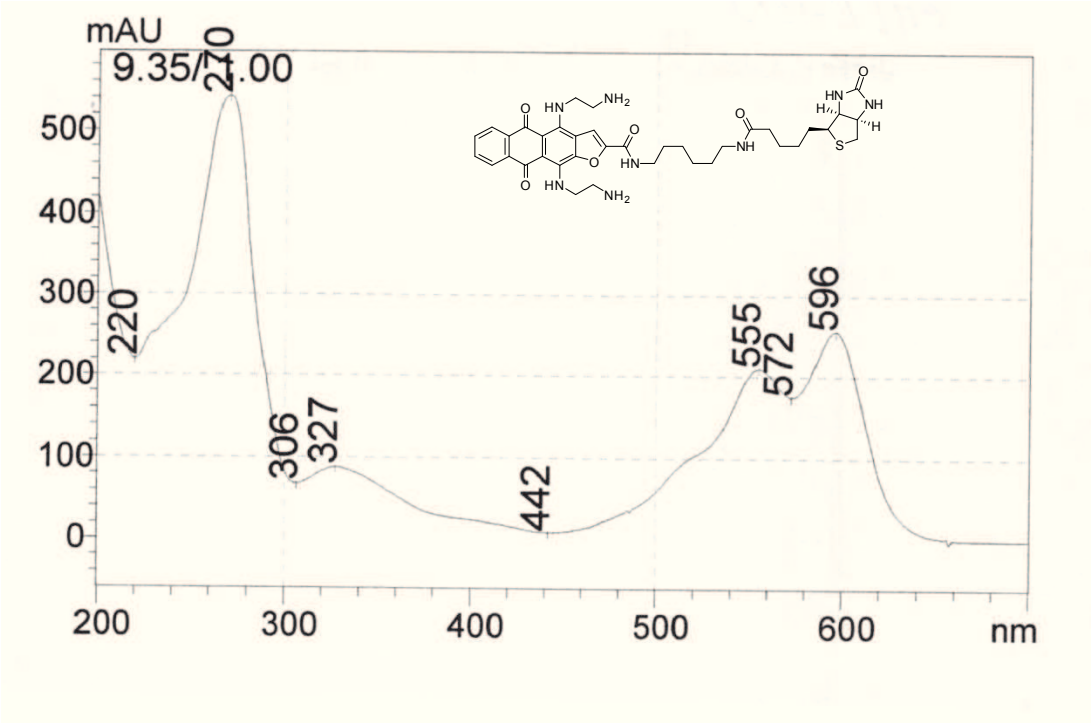

DMSO-d6
182.05
181.65
171.87
162.75
156.92
151.61
150.47
142.60
133.99
133.80
133.64
127.30
126.10
126.00
123.33
120.43
108.44
61.91
61.71
61.68
61.07
59.22
55.47
38.31
35.25
29.15
28.24
28.06
25.14
25.36
Chemical Shift (ppm)

*4,11-bis((2-Aminoethyl)amino)-biotinyl-N-(6-aminohexyl)-5,10-dioxo-5,10-dihydroanthra[2,3-b]furan-2-carboxamide (**6**, **b-2a**)*



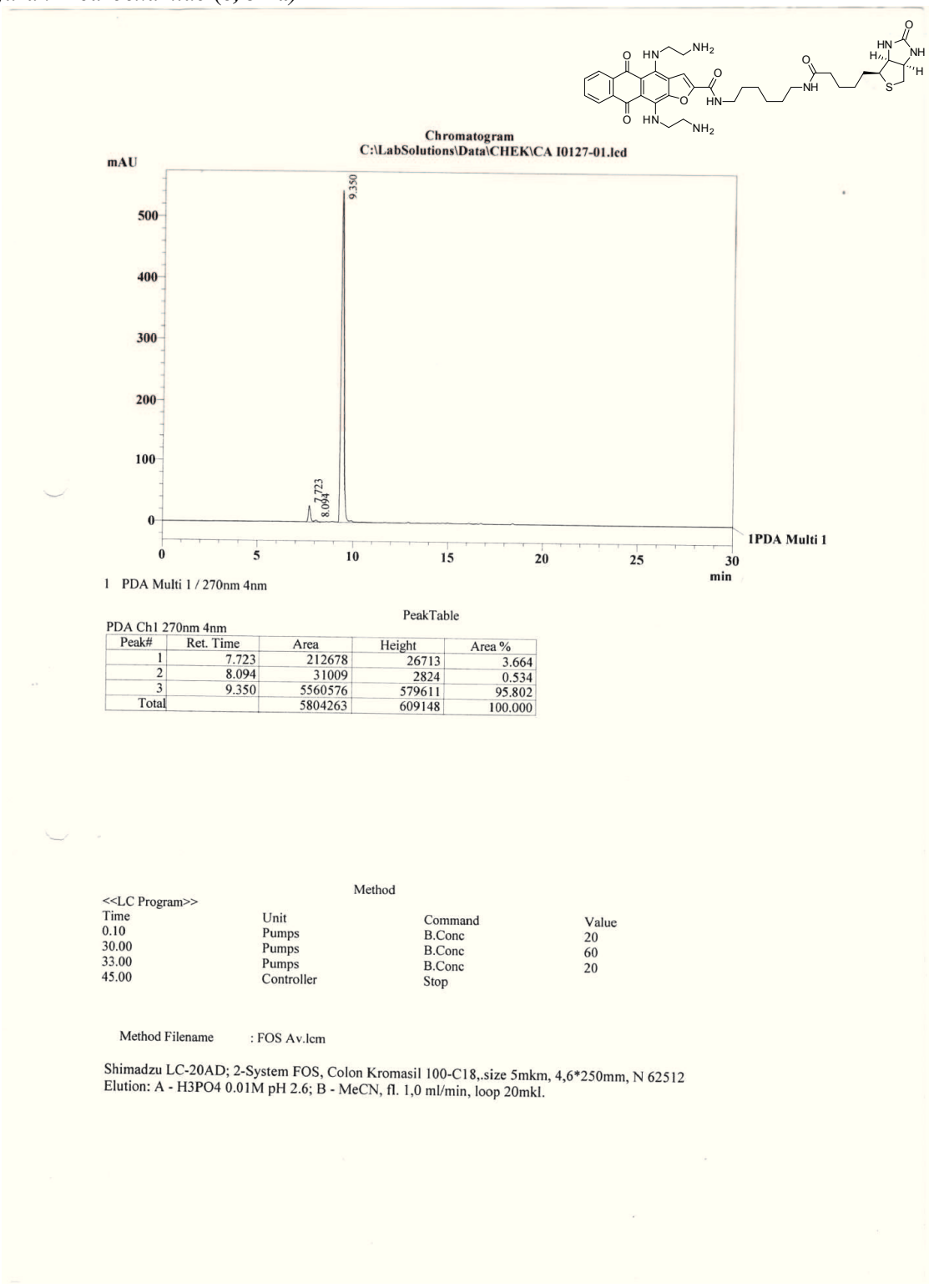
**S8.2. UV-Spectra**

*4,11-bis((2-Aminoethyl)amino)-biotinyl-N-(6-aminohexyl)-5,10-dioxo-5,10-dihydroanthra[2,3-b]furan-2-carboxamide (**6, b-2a**)*



## S8.3. HPLC-tracks

*4,11-bis((2-Aminoethyl)amino)-biotinyl-N-(6-aminohexyl)-5,10-dioxo-5,10-dihydroanthra[2,3-b]furan-2-carboxamide* (***6, b-2a***)




1 PDA Multi 1 / 270nm 4nm

PeakTable

PDA Ch1 270nm 4nm

| Peak# | Ret. Time | Area | Height | Area % |
|---|---|---|---|---|
| 1 | 7.723 | 212678 | 26713 | 3.664 |
| 2 | 8.094 | 31009 | 2824 | 0.534 |
| 3 | 9.350 | 5560576 | 579611 | 95.802 |
| Total | | 5804263 | 609148 | 100.000 |

Method

<<LC Program>>

| Time | Unit | Command | Value |
|---|---|---|---|
| 0.10 | Pumps | B.Conc | 20 |
| 30.00 | Pumps | B.Conc | 60 |
| 33.00 | Pumps | B.Conc | 20 |
| 45.00 | Controller | Stop | |

Method Filename : FOS Av.lcm

Shimadzu LC-20AD; 2-System FOS, Colon Kromasil 100-C18,.size 5mkm, 4,6*250mm, N 62512
Elution: A - H3PO4 0.01M pH 2.6; B - MeCN, fl. 1,0 ml/min, loop 20mkl.

**Section D:**

# Potent anti-Kirsten ras (KRAS) single-stranded miRNA oligonucleotide mimics as therapeutic agents.

G. Miglietta, A. Ferino, S. Cogoi, R. Picco, S. Vogel, J. Wengel and LE. Xodo

(in preparation)

# Potent anti Kirsten ras (*KRAS*) single-stranded miRNA oligonucleotide mimics as therapeutic agents

Giulia Miglietta,[a] Annalisa Ferino,[a] Susanna Cogoi,[a] Raffaella Picco,[a] Stefan Vogel,[b] Jesper Wengel [b] and Luigi E. Xodo [a] *

(a) Department of Medicine, Laboratory of Biochemistry, University of Udine, Italy; **luigi.xodo@uniud.it;**

(b) Nucleic Acids Centre, University of Southern Denmark, DK 5230, Odense, Denmark

**Corresponding author**: Luigi E. Xodo, Department of Medicine, Laboratory of Biochemistry, University of Udine, Italy; **luigi.xodo@uniud.it**;

**Keywords**: KRAS, PDAC cells, KRAS, miR-216b, POPC liposome

**ABSTRACT**

**Datasets reporting microRNA expression profiles in normal and cancer cells showed that miR-216b is aberrantly downregulated in pancreatic ductal adenocarcinoma (PDAC). Oncogenic *KRAS*, which drives the pathogenesis of PDAC, is a target of miR-216b. We designed single-stranded miR-216b mimics, engineered with unlocked nucleic acid (UNA) modifications, and found that they strongly suppress *KRAS* in PDAC cells. We also report a new delivery strategy for miR-216b, which is based on the use of palmitoyl-oleyl-phosphatidylcholine (POPC) liposomes functionalized with lipid-modified miR-216b and lipid-modified cell penetrating TAT peptide.**

## 1. Introduction

Pancreatic ductal adenocarcinoma (PDAC) is one of the major causes of death in western countries **[1, 2].** As current treatments are not effective, there is an urgent need to develop new therapies **[3]**. The main genetic lesion present in > 90 % of PDAC patients is a mutation in the *KRAS* proto-oncogene, mainly in exon 1 at codon 12, G12D (Gly→Asp) or G12C (Gly→Cys) **[4]**. It has been demonstrated that mutant *KRAS* is the major driver of PDAC **[4-6]** and that the expression of *KRAS* $^{G12D}$ in the pancreas of transgenic mice causes intraepithelial neoplasia lesions that progress into full malignancy **[6]**. The expression of *KRAS* $^{G12D}$ is necessary for tumour maintenance and its extinction leads to a rapid tumour regression **[6]**. Pancreatic cancer cells are indeed "addicted" to mutant *KRAS,* as this oncogene reprograms the glucose and glutamine metabolism in the tumour cells **[7, 8]**.

Recent studies have shown that several functions in cancer cells, including proliferation, migration, invasion and gene expression, are regulated by small noncoding microRNAs **[9-14].** They are synthesised as long RNA strands that fold into hairpin-loop structures processed by Drosha and Dicer into mature duplexes ranging from 17 to 26 nt in length **[9-13]**. The guide strand of the mature RNA duplexes forms a complex with Ago proteins that binds to a 3'-UTR mRNA target. This mediates two modes of gene silencing: translation repression and/or RNA decay **[12, 13]**. Synthetic double-stranded (ds) miRNAs gave encouraging results as

antigene molecules **[15]** and recent studies have demonstrated that single-stranded (ss) RNAs, mimicking the guide strand of miRNAs, can also mediate an Ago-dependent inhibition of the target gene **[16-21]**. These findings provided new perspectives on the use of synthetic miRNAs as therapeutic agents **[17]**.

We designed single-stranded miRNA mimics to inhibit *KRAS* in pancreatic cancer cells. We started from the observation that certain small noncoding RNAs are aberrantly expressed in cancer tissues **[23-27]**. It is now common knowledge that miRNAs inhibiting the expression of tumour suppressor genes are oncogenic in nature and favour the development of cancer **[22-27]**. In contrast, miRNAs down-regulating the expression of oncogenes behave as tumour suppressors and inhibit cancer growth **[22-27]**. In this study, we focused on a miRNA aberrantly down-regulated in PDAC, miR-216b, with the aim of obtaining a therapeutic agent capable to potently suppress *KRAS* in PDAC cells. We found that single-stranded miR-216b with one or two unlocked nucleic acid modifications, strongly suppress oncogenic *KRAS* in PDAC cells. We also report data regarding the activity of miR-216b conjugated to a palmityl chain and fixed on the surface of palmitoyl-oleyl-phosphatidylcholine (POPC) liposomes functionalized with the trans-activator of transcription of the human immune-deficiency virus (TAT) cell penetrating peptide.

## 2. Results and discussion

We started our study from the observation that abnormally high levels of oncogenic miRNAs and low levels of tumour suppressor miRNAs have been found in PDAC **[27, 28]**. To identify miRNAs specific for oncogenic *KRAS*, we compared three datasets reporting miRNA expression profiles in normal and PDAC cells: E-MTAB-753, GSE43796 and GSE41372 **[29]**. Among the abnormally downregulated miRNAs we focused on miR-216b (from now on 216b), which showed an expression fold change from 2.27 to 27.95 (**Supplementary Data, S1**). A direct link between 216b and *KRAS* has been found in nasopharyngeal tumours **[30]**. Additionally, in a transgenic mouse model, 216b was found to be downregulated during PDAC development, confirming its tumour suppressor nature **[31]**. As single-stranded (ss) siRNAs act through the RNAi pathway and silence gene expression as the duplex analogues do **[20, 21]**, ss-miRNAs might also be bioactive. Indeed, ss-miRNAs are

also loaded into the RNA-induced silencing complex and inhibit gene expression **[16-19]**. In the light of these studies, we designed synthetic ss-miRNA to suppress *KRAS* in pancreatic cancer cells. One might wonder why using ss-miRNAs, considering that ds-miRNAs are potent tools for gene silencing. There are mainly two reasons: (i) ds-miRNAs, being more complex than ss-miRNAs, are not so efficiently transported into the cells as their single-stranded analogues **[32];** (ii) ds-miRNAs are composed by the seed and passenger strands, and the latter may cause undesired off-target effects **[33]**. **Figure 1A** shows the sequence of 216b and its *KRAS* target in the 3'-untranslated region (3'-UTR) of mRNA. The sequence at the 5'-end of 216b, UCUCUAAA-5', is perfectly complementary to the mRNA target and represents the "seed region". Some complementarity with the target is also present at the 3' half of 216b: 5'-UAAAC is base-paired with mRNA. This should improve the interactions between the miRNA and the target gene **[34, 35]**. In order to increase the nuclease stability, we designed 216b mimics with one or two unlocked nucleic acid (UNA) modifications **[36, 37]**. The key feature of UNA is the loss of C2'-C3' bond of the ribose, a modification that makes the strand more flexible than the native oligonucleotide (**Figure 1B**) **[38]**. Previous studies have shown that a single UNA modification in the middle of a RNA/RNA duplex can lower the $T_m$ by 5–10 °C, but when the UNA-modified base is placed near the duplex end, it causes a rather modest drop in $T_m$ of 1–3 °C **[38, 39]**.

We designed two UNA-modified 216b mimics: U1 with one modified adenine (uA) at the 3'-end; U2 with two uAs, at the 3'-end and in the middle of the oligoribonucleotide (**Figure 1**, **Supplementary Table S2, Table S3**). The uA incorporated in the middle of the sequence is located outside the seed sequence. The uA at the 3'-end is also located in a portion of 216b that does not pair with mRNA. The UNA modifications introduced in 216b should therefore not affect the hybridization of the UNA-modified mimics to the RNA target. Previous studies have reported that (i) ss-miRNAs with 2'-methyl, 2'-fluoro, 2'-O-methoxyethyl and with a phosphorothioate backbone are fairly active **[16-19]**; (ii) the binding of Ago2 to modified siRNAs (including UNA-modified siRNAs) is not affected by the chemical modifications **[18]**. We introduced UNA modifications in our mimics, as they can be considered less "invasive" than those mentioned above. Chorn et al.

compared the activity of fluorinated and 2'-O-methyl modified ss-miRNAs with and without a 5'-phosphate, and found that phosphorylation increases the activity, although not in all constructs tested **[16-19].** We therefore designed UNA-modified ss-miRNAs with and without a phosphate group at the 5'-end (**Figure 1A**). To investigate the gain in stability given by the UNA insertions, we subjected the wild-type and UNA-modified mimics to the action of cellular nucleases of a total cell extract. **Figure 2A** shows the integrity of ss-miRNAs with a phosphate at the 5' end after incubation in a total Panc-1 extract for 1, 2 and 4 h, at 37 °C. It can be seen that after 4 h, only ~ 10 % of 216b-P is intact. This percentage increases to ~ 20 and ~ 65 %, respectively, when one or two UNA modifications are introduced in the oligonucleotide. **Figure 2B** shows that the non-phosphorylated analogues are slightly more resistant: after 4 h, ~ 27 % of 216b is intact, while the percentage of undigested U1 and U2 is ~ 30 and ~80 %, respectively. These data clearly show that the insertion of UNA at the 3'-end and in the middle of the sequence boosts the resistance of the designed mimics to the cellular nucleases in PDAC cells. There is only one study in the literature reporting that a replacement strategy with 216b is effective in down-regulating oncogenic *KRAS*, but this was found in nasopharyngeal cancer cells. Although 216b is aberrantly downregulated in PDAC cells, its therapeutic potential in pancreatic cancer cells has not been investigated so far. Considering that no effective anti-KRAS therapeutic agents have been found up to now, to investigate the potency of 216b in PDAC is an important issue. Against this background, our first step was to demonstrate that 216b is indeed *KRAS* specific in PDAC cells. For this purpose, we hybridized wild-type and UNA-modified 216b miRNAs to the complementary sequence to obtain 22-mer ds-mimics with or without UNA insertions, as well as a mutated ds-UNA-modified miRNA. Panc-1 cells were transfected two times with the duplexes and after an incubation of 48 h, a Western blot was performed to measure the level of protein KRAS. Duplexes ds-U1 and ds-U2 dramatically reduced the level of protein KRAS to ~ 10 % of the control (mutated duplex) (**Figure 3**). Instead, the wild-type duplex produced a weaker effect on protein expression (~ 35 % of the control). This experiment clearly demonstrates that in PDAC cells oncogenic *KRAS* is a target of miR-216b.

Next, we examined the capacity of the designed ss-miRNA mimics, with or without a phosphate at the 5' end, to suppress *KRAS* in Panc-1 cells. As miRNAs inhibit gene expression by translation repression and/or mRNA decay, **[12-14]** we determined by quantitative RT-PCR the level of the *KRAS* transcript in Panc-1 cells treated with the various forms of 216b. As observed in the nasopharyngeal cells, **[30]** we did not notice any significant reduction of *KRAS* mRNA at 16, 40 or 72 h (**Supplementary Data, S4**). Instead, when we measured by Western blot the level of the KRAS protein in Panc-1 cells treated with the various ss-miRNA mimics, we observed a dramatic suppression of the protein. In keeping with the nuclease stability data, the compound showing the highest capacity to suppress protein KRAS was the miRNA with two UNA modifications. **Figure 4A, B** shows that both phosphorylated and non-phosphorylated mimics caused a strong suppression of the protein. The UNA modifications turned out to be essential for the activity of the mimics, which brought the protein down to < 10 % of the control, while the 5' phosphorylation is not essential, at least for this specific miRNA **[16]**. Note that the mutated sequence (mut) did not impact *KRAS* expression. As a next issue, we wondered what the specificity of 216b for *KRAS* is, since the oncogene is highly homologous to *HRAS* and *NRAS*. As illustrated in **Figure 4C, D**, none of the designed 216b mimics had effect on HRAS and NRAS proteins. This was expected, as the matching between the seed region of the designed mimics and the 3'-UTR sequence in these genes is suboptimal. In these cases, the base pairing in the miRNA 3' half would assume importance for stabilizing the interaction **[12-14]**. The lack of complementarity at the 3' end makes the mimics ineffective against *HRAS* and *NRAS*. We can therefore conclude that, at least in pancreatic cancer cells, 216b regulates only one gene of the ras family.

To obtain further experimental evidence that the designed UNA-modified ss-miRNA mimics are active, we tested them in a luciferase assay. We used plasmid pGL4.75-KRAS LCS6m, carrying *Renilla* luciferase driven by the CMV promoter and the human *KRAS* 3'-UTR (3200 bp) containing the 216b target. Panc-1 cells were treated with 10 nM mimics complexed with *Interferin* (Polyplus, France) and with 30 ng pGL4.75-KRAS LCS6m and 70 ng pGL3-Control Vector carrying *Firefly*

luciferase driven by the SV40 promoter. The cells were treated twice with single-stranded 216b, U1 and U2: one treatment 24 h after the other. The three ss-miRNAs were able to reduce luciferase to an extent that correlates with the modifications introduced in 216b (**Figure 5**). The molecule with the strongest anti-*KRAS* activity was U2, which decreased luciferase to ~ 40 % of the control (luciferase expressed in mut-treated Panc-1 cells). A roughly similar result was obtained with the 5' phosphorylated mimics (not shown). The luciferase experiment unambiguously proved that oncogenic *KRAS* is downregulated by 216b. The luciferase assay gave a weaker effect compared to the western blot assay probably because: (i) CMV is a promoter stronger than the *KRAS* promoter; (ii) the luciferase assay is based on a non-natural target.

Finally, since *KRAS* stimulates the survival and proliferation pathways, the suppression of protein KRAS should result in the inhibition of colony formation. Indeed we found that in Panc-1 and MIA PaCa-2 pancreatic cancer cells the number of colonies was reduced ~ 50 % by the designed miRNAs.

Since *KRAS* stimulates the pathway controlling cell growth, the suppression of protein KRAS should result in the inhibition of proliferation. To assess the effect of 216b, U1 and U2 on the growth of PDAC cells we carried out colony formation assays with Panc-1 and MIA PaCa-2 cells, carrying the *KRAS* mutations G12D and G12V, respectively. The PDAC cells were seeded in a medium after being diluted in a way that a single colony could be formed from each cell. After ~10 days of growth, the colonies of at least 50 cells were counted and the results plotted in a histogram. **Figure 6** shows the results of a typical colony-formation assay obtained with Panc-1 and MIA PaCa-2 cells. Note that the number of colonies in the untreated plate is similar to that measured on the plate treated with the control mut, which does not suppress oncogenic *KRAS*. In contrast, 216b and its UNA-analogues strongly reduced the number of colonies from both types of cells. U2 reduced the number of MIA PaCa-2 colonies by ~ 60 %. As *KRAS* controls cell adhesion via the integrin-linked kinase (ILK) [**40-42**] its suppression in Panc-1 cells resulted in some cell aggregation. The number of colonies was reduced by ~ 50 %.

In the experiments described above, the miRNAs were delivered to the cells as complexes with *Interferin* (Polyplus, France), a commercial transfecting agent. To improve the delivery of miR-216b, we used a transport system based on the low toxicity of palmitoyl-oleyl-phosphatidylcholine (POPC) liposomes in combination with surface attached functionalities [**43**]. The POPC liposomes were functionalized with the trans-activator of transcription of the human immune-deficiency virus (TAT) as a cell-penetrating peptide and with 216b, as previously reported [**44, 45**]. Both 216b and TAT were chemically modified with a palmitoyl membrane anchor to allow their rapid attachment to the liposome surface [**46**]. The strategy is illustrated in **Figure 7A**. The POPC liposomes were treated with both lipid-modified 216b and TAT, which spontaneously anchored to the liposome surface. Oligonucleotide 216b being not covalently attached to the liposomes, can move freely on the lipid surface and interact efficiently with the mRNA target. The membrane anchor of miR-216b consists of a 3-amino-1,2-propanediol unit with two saturated palmityl chains (membrane anchor Y) [**46**]. We treated two times the PDAC cells Panc-1, MIA PaCa-2 with the POPC liposomes functionalized with TAT and miR-216b and measured by Western blot the level of protein KRAS (**Figure 7B**). It can be seen that KRAS protein is reduced to about half the level present in untreated cells or cells treated with a mutant miRNA. This result suggests that POPC liposome are an attractive vehicle to deliver miR-26b in vivo. Finally, the bioactivity of miR-216b delivered via POPC was tested in a colony formation assay. **Figure 8** shows that the miR-216b fixed on POPC reduce Panc-1 colony formation by about 50%.

## 3. Conclusion

In conclusion, our results show that 216b, which is aberrantly downregulated in PDAC, is a microRNA that targets the *KRAS* oncogene. When UNA modifications are introduced outside the seed region, single-stranded 216b, with and without a 5'phosphate, show a strong activity and promotes a dramatic suppression of the KRAS protein. This results in the inhibition of colony formation by PDAC cells. We have shown for the first time that 216b behaves as a tumour suppressor miRNA in PDAC cells. The UNA modifications make 216b stable against the cellular nucleases and quite a strong therapeutic agent for PDAC. In addition, we

conjugated miR-216b to a palmityl chain and fixed the effector oligonucleotide to the surface of POPC liposome, which were functionalized with TAT peptide. The results obtained with three PADC cell lines showed that the nanoparticle functionalized with TAT and miR-216b cause a 50 % decrease of protein KRAS and a 40% inhibition of colony formation. Our data suggest that this delivery strategy will be helpful for *in viv*o studies.

## 4. Experimental part

### 4.1 General

Detailed procedures for cell culture transfections, dual luciferase assay, Western blots, RT-PCR, oligonucleotide stability in cell environment are given in the Supplementary material.

### 4.2 Synthesis of UNA-modified miRNAs.

The synthesis of the UNA-modified miRNA 216b mimics (216b, U1, U2, mut, U1-P, U2-P) was performed on an automated nucleic acid synthesizer as previously reported [**43**]. The oligonucleotides were purified by HPLC by using a C18 3 µm 300 Å reversed phase column as previously described [**46**]. Maldi MS analyses were carryout out to confirm the structure of the designed UNA-modified miRNAs and evaluate their purity (See Supplementary). Calculated and experimental masses were almost identical for all the oligonucleotides (Table 2) and Supplementary data.

### 4.3 Synthesis of lipid-conjugated TAT and miR-216b

The oligonucleotide was synthesized on an ExpediteTM 8900 nucleic acid synthesis system (Perceptive Biosystems Inc.). The synthesis was performed on a 1.0 µmol scale on GE Healthcare Custom Primer Support$^{TM}$ T40s using standard conditions for automated synthesis with DCI as activator. However, the lipid modified phosphor amidite was dissolved in 2:1 DCE: MeCN at a concentration of 0.1 M, 42$^{®}$ was used as activator (Proligo reagent/Sigma-Aldrich) and the coupling time was increased to 20 min. The DMT protecting group on the last nucleotide in the sequence was removed. After deprotection and cleavage from the solid support using standard conditions (conc. $NH_3$ (aq.) over night at 55 ºC), the oligonucleotides were purified by HPLC.

TAT was synthesized on a Liberty 1 Microwave Peptide Synthesizer using a Rink Amide resin and an amino acid concentration of 0.2 M for unmodified amino acids and 0.1 M for the lipid-modified amino acid. After synthesis, the peptide was cleaved from the resin, deprotected and purified by HPLC using a C18 5 μm 10 x150 mm column, as previously described

**4.4 Colony forming assay.**

Since *KRAS* stimulates the pathway controlling cell growth, the suppression of protein KRAS should result in the inhibition of proliferation. To assess the effect of 216b, U1 and U2 on the growth of PDAC cells we carried out colony formation assays with Panc-1 and MIA PaCa-2 cells, carrying the *KRAS* mutations G12D and G12V, respectively. The PDAC cells were seeded in a medium after being diluted in a way that a single colony could be formed from each cell. After 7-13 days of growth, the colonies of at least 50 cells were counted and the results plotted in a histogram. The figure below shows the results of a typical colony-formation assay obtained with Panc-1 and MIA PaCa-2 cells. Note that the number of colonies in the untreated plate is similar to that measured on the plate treated with the control mut, which does not suppress oncogenic *KRAS*. In contrast, 216b and its UNA-analogues strongly reduced the number of colonies from both types of cells. U2 reduced the number of MIA PaCa-2 colonies by ~ 60 %. As *KRAS* controls cell adhesion via the integrin-linked kinase (ILK) **[40-42]** its suppression in Panc-1 cells resulted in some cell aggregation. The number of colonies was reduced by ~ 50 %. MiRNA 216b and the UNA-modified mimics U1 and U2 decrease the % of colonies in Panc-1 and MIA PaCa-2 cancer cells. NT= untreated cells. Histogram shows the % colonies in miRNA-treated cells compared to NT. Values obtained from three independent experiments. *=P < 0.05.

**Cell Culture and transfections**

Human pancreatic adenocarcinoma Panc-1 and MIA PaCa-2 cells were maintained in exponential growth in Dulbecco's Modified Eagle's Medium (DMEME) containing 100 U/mL penicillin, 100 mg/mL streptomycin, 20 mM L-glutamine and 10 % fetal bovine serum (Euroclone, Italy). The cell lines have been genotyped by Microsynth (Switzerland) and found that they matched 100% with the DNA-profiles of Panc-1 (ATCC® CRL-1469TM)

and MIA PaCa-2. The cells were transfected with the various miRNAs twice after 24 h of incubation in a humidified atmosphere containing 5 % $CO_2$ at 37 °C. The expression vectors were transfected with *jet PEI* (Polyplus, France) while the miRNas were transfected with Interferin (Polyplus, France) following manufacturer's instructions. Double-stranded miRNA mimics have been prepared in 100 mM NaCl and 50 mM Tris-HCl pH 7.4, heated 5 min at 80 °C and annealed overnight at room temperature.

**Stability of miRNAs in cellular environment**

To determine their nuclease resistance, the designed miRNAs (3 μM) have been incubated for 1, 2 and 4 h at 37°C in a total nuclear extract from Panc-1 cancer cells (2 μg) obtained as previously described (Nucleic Acids Res. 2008, 36, 3765-80). After incubation the oligonucleotides have been run in a denaturing 20 % polyacrylamide gel (7 M urea, 1 x TBE), which was stained with "stains all".

**Dual luciferase assays**

Panc-1 cells were plated (15 x $10^3$) in 96-well plate and after one day transfected with miRNA 216b or UNA-analogues and with plasmids pGL4.75-KRAS LCS6m (a gift from Frank Slack, Plasmid # 44571, Addgene) and pGL3 Control Vector (Promega, USA). Transfection was performed by mixing 70 ng of pGL3 Control Vector (*Firefly luciferase)* with 30 ng of pGL4.75-KRAS LCS6m (*Renilla luciferase)*, by using jet-PEI (Polyplus) as a transfectant agent. A second miRNA transfection was performed 24 h after the first transfection and the luciferase assays were performed 48 h after the second transfection. A Dual-Glo Luciferase Assay System (Promega, USA) was used. Samples were read on a Turner Luminometer and the relative luminescence expressed as (T/C x 100) where T= *Renilla* luciferase/*Firefly* luciferase in miRNA treated cells and C= *Renilla* luciferase/*Firefly* luciferase in mut-treated cells.

**Western blots**

Total protein lysates (10-15 μg) extracted from Panc-1 cells were sonicated for 10 minutes and the lysates were electrophoresed on 12 % SDS-PAGE and transferred into a nitrocellulose membrane, at 70 V for 2 h. The filter was blocked for 1 h with 5 % nonfat dry milk solution in PBS 0.05 % Tween (Sigma-Aldrich, Italy) at room temperature.

Membranes were incubated overnight at 4 °C with the primary antibodies: monoclonal anti-KRAS (clone 3B10-2F2, IgG1 mouse, 2.5 µg/mL, Sigma-Aldrich, USA), polyclonal anti-HRAS (IgG rabbit, diluted 1:300, Santa Cruz Biotechnology Inc, USA), monoclonal anti-NRAS (clone F155-227, $IgG_1$ mouse, 2.5 µg/mL) and monoclonal anti-actin (clone JLA20, IgM mouse, $1x10^{-4}$ µg/mL, Calbiochem, Merck Millipore, Germany). The membranes were washed with a 0.05 % Tween in PBS and then incubated 1 h with the secondary antibodies horseradish peroxidase conjugated: anti-mouse IgG (diluted 1:5000) and anti-mouse IgM (diluted 1:2000) (Calbiochem, Merck Millipore, Germany) and anti-rabbit IgG (diluted 1:5000) (Calbiochem, Merck Millipore, Germany). To detect the protein we used Super signal ® West PICO and FEMTO (ThermoFisher Scientific Pierce, USA). The exposure time depended on the antibody used and was usually between 30 s and 5 min. The protein levels were quantified by the Image Quant TL Version 2003 software (Amersham).

**Acknowledgment**

This work has been carried out with the financial support of Italian Association for Cancer Research, P.C.14301, 2015.

**References**

1 A. Jemal, R. Siegel, E. Ward, T. Murray, J. Xu, M.J. Thun, Cancer statistics, 2007, Cancer J. Clin. 57 (2007) 43-66.

2 A. Jemal, R. Siegel, J. Xu, E. Ward, Cancer statistics 2010, Cancer J. Clin. 60 (2010) 277-300.

3 Kleeff, C. Reiser, U. Hinz, J. Bachmann, J. Debus et al, Surgery for recurrent pancreatic ductal adenocarcinoma, Ann. Surg. 245 (2007) 566-572.

4 K.L. Bryant, J.D. Mancias, A.C. Kimmelman, C.J. Der, KRAS: feeding pancreatic cancer proliferation, Trends Biochem. Sci. 39 (2014) 91-100.

5 M.A. Collins, F. Bednar, Y. Zhang, JC. Brisset, CJ. Galbàn, S. Rakshit, KS. Flannagan, NV. Adsay, MP Di Magliano, Oncogenic Kras is required for both the initiation and maintenance of pancreatic cancer in mice, J. Clin. Invest. 122

(2012) 639-653.

**6** B. Kong, C.W. Michalski, M. Erkan, H. Friess, J. Kleeff, From tissue turnover to the cell of origin for pancreatic cancer, Nat. Rev. Gastroenerol. Hepatol. 8 (2011) 467-472.

**7** H. Ying, A.C. Kimmelman, C.A. Lyssiotis, S. Hua, G.C. Chu, E. Fletcher-Sananikone, JW. Locasale, J. Son, H. Zhang, Oncogenic Kras Maintains Pancreatic Tumors through Regulation of Anabolic Glucose Metabolism, Cell 149 (2012) 656-670.

**8** J. Son, C.A. Lyssiotis, H. Ying, X. Wang, S. Hua, M. Ligorio, RM. Perera, CR. Ferrone, E. Mullarky, N. Shyh-Chang, Y. Kang, JB. Fleming, N. Bardeesy, JM. Asara, MC. Haigis, RA. De Pinho, LC. Cantley, AC. Kimmelman, Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway, Nature 496 (2013) 101-105.

**9** S. Jonas, E. Izaurralde, Towards a molecular understanding of microRNA-mediated gene silencing, Nat. Rev. Genet. 16 (2015) 421-433.

**10** G. Stefani, F.J. Slack, Small non-coding RNAs in animal development, Nat. Rev. Mol. Cell Biol. 9 (2008) 219-230.

**11** J. Krol, I. Loedige, W. Filipowicz, The widespread regulation of microRNA biogenesis, function and decay, Nat. Rev. Genet. 11 (2010) 597-610.

**12** W. Filipowicz, S.N. Bhattacharyya, N. Sonenberg, Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?, Nat. Rev.Genet. 9 (2008) 102-114.

**13** E. Huntzinger, E. Izaurralde, Gene silencing by microRNAs: contributions of translational repression and mRNA decay, Nat. Rev. Genet. 12 (2011) 99-110.

**14** H.O. Iwakawa, Y. Tomari, The Functions of MicroRNAs: mRNA Decay and Translational RepressionTrends, Cell Biol. 25 (2015) 651-665.

**15** M.A. Goldgraben, R. Russell, O.M. Rueda, C. Caldas, A. Git, Double-stranded microRNA mimics can induce length- and passenger strand–dependent effects in a cell type–specific manner, RNA 22 (2016) 193-203.

**16** G. Chorn, M. Klein-McDowell, L. Zhao, M.A. Saunders, W.M. Flanagan, A.T. Willingham, L.P. Lim, Single-stranded microRNA mimics, RNA 18 (2012) 1796-1804.

**17** M. Matsui, T.P. Prakash, D.R. Corey, Argonaute 2-dependent Regulation of Gene Expression by Single-stranded miRNA Mimics, Mol. Ther. 24 (2016) 946-955.

**18** N.T. Schirle, G.A. Kinberger, H.F. Murray, W.F. Lima, T.P. Prakash, I.J. MacRae, Structural analysis of human Argonaute-2 bound to a modified siR-NA guide, J. Am. Chem. Soc. 138 (2016) 8694-8697.

**19** JK. Watts, D. Yu, K. Charisse, C. Montaillier, P. Potier, M. Manoharam, D.R. Corey, Effect of chemical modifications on modulation of gene expression by duplex antigene RNAs that are complementary to non-coding transcripts at gene promoters, Nucleic Acids Res. 38 (2010) 5242-5259.

**20** D. Yu, H. Pendergraff, J. Liu, H.B. Kordasiewicz, D.W. Cleveland, EE. Swayze, WF. Lima, ST. Crooke, TP. Prakash, DR. Corey, Single-Stranded RNAs Use RNAi to Potently and Allele-Selectively Inhibit Mutant Huntingtin Expression, Cell 150 (2012) 895-908.

**21** WF. Lima, TP. Prakash, HM. Murray, GA. Kinberger, W. Li, GA. Kinberger, W. Li. AE.Chappell, CS. Li, SF.Murray, H. Gaus, PP.Seth, EE. Swayze, ST. Crooke, Single-stranded siRNAs activate RNAi in animals, Cell 150 (2012) 883-894.

**22** M. Matsui, D. Corey, Non-coding RNAs as drug targets, Nat. Rev. Drug Discov. 16 (2017) 167-179.

**23** S. Lin, R.I. Gregory, MicroRNA biogenesis pathways in cancer, Nat. Rev. Cancer 15 (2015) 321-333.

**24** C.L. Bartels, G.J. Tsongalis, MicroRNAs: Novel Biomarkers for Human Cancer Clin. Chem. 55 (2009) 623-631.

**25** H. Ling, M. Fabbri, G.A. Calin, MicroRNAs and other non-coding RNAs as targets for anticancer drug development, Nat. Rev. Drug Discov. 12 (2013) 847-865.

**26** G. Di Leva, C. Croce, miRNA profiling of cancer, Curr. Opin. Genet. Dev. 23 (2013) 3-11.

**27** M. Seux, J. Iovanna, J.C. Dagorn, N.J. Dusetti, MicroRNAs in Pancreatic Ductal Adenocarcinoma: New Diagnostic and Therapeutic Clues, Pancreatology 9 (2009) 66-72.

**28** M. Diab, I. Muqbil, R.M. Mohammad, A.S. Azmi, P.A. Phillip, MiR-142 inhibits

the development of cervical cancer by targeting HMGB1, J. Clin. Med. 8 (2016) 4001-4007.

**29** T. Barrett, DB. Troup, SE. Wilhite, P. Ledoux, C. Evangelista, IF. Kim, M. Tomashevsky, KA. Marshall, KH. Phillippy, PM. Sherman, RN. Muertter, M.Holko, O. Ayanbule, A. Yefanov, A. Soboleva, NCBI GEO: archive for functional genomics data sets—10 years on, Nucleic Acids Res. 39 (2011) 1005–1010.

**30** M. Deng, H. Tang, Y. Zhou, M. Zhou, W. Xiong, Y. Zheng, Q. Ye, X. Zeng, Q. Liao, X. Guo, X. Li, J. Ma, G. Li, miR-216b suppresses tumor growth and invasion by targeting KRAS in nasopharyngeal carcinoma, J. Cell Sci. 124 (2011) 2997-3005.

**31** AC. Azevedo-Pouly, DS. Sutaria, J. Jiang, OA. Elgamel, F. Amari, D. Allard, PJ. Grippo, V. Coppola, TD. Shmittgen, miR-216 and miR-217 expression is reduced in transgenic mouse models of pancreatic adenocarcinoma, knockout of miR-216/miR-217 host gene is embryonic lethal, Funct. Integr. Genomics 17 (2017) 203-212.

**32** E. Koller, T.M. Vincent, A. Chappell, S. De, M. Manoharan, C.F. Bennett, Mechanisms of single-stranded phosphorothioate modified antisense oligonucleotide accumulation in hepatocytes, Nucleic Acids Res. 39 (2011) 4795-4807.

**33** AL. Jackson, SR. Bartz, J. Schelter, SV. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet, PS. Linsley, Expression profiling reveals off-target gene regulation by RNAi, Nat. Biotecnol. 21 (2003) 635-637.

**34** J. Brennecke, A. Stark, RB. Russell, SM. Cohen, Principles of MicroRNA–Target Recognition, Plos Biol. 2005, 3, e85.

**35** A. Grimson, K.K. Farh, W.K. Johnston, P. Garrett-Engele, L.P. Lim and D.P. Bartel, MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing Mol. Cell 27 (2007) 91-105.

**36** JB. Bramsen, MB. Laursen, AF. Nielsen, TB. Hansen, C. Bus, T. Højland, M. Abramov, A. Van Aerschot, D. Odadzic, R. Smicius, J. Haas, C. Andree, J. Barman, M. Wenska, P. Srivastava, C. Zhou, D. Honcharenko, S. Hess, E. Müller, GV. Bobkov, SN. Mikhailov, E. Fava, TF. Meyer, J. Chattopadhyaya, M.

Zerial, JW. Engels, P. Herdewijn, J. Wengel, J. Kjems, A large-scale chemical modification screen identifies design rules to generate siRNAs with high activity, high stability and low toxicity, Nucleic Acids Res. 37 (2009) 2867-2881.

**37** NM. Snead, JR Escamilla-Powers, JJ. Rossi, AP. McCaffrey, 5′ Unlocked Nucleic Acid Modification Improves siRNA Targeting, Mol. Ther. Nucleic Acids, 2 (2013) e103.

**38** N. Langkjaer, A. Pasternak, J. Wengel, UNA (unlocked nucleic acid): A flexible RNA mimic that allows engineering of nucleic acid duplex stability, Bioorg. Med. Chem. 17 (2009) 5420-5425.

**39** A. Pasternak, J. Wengel, Unlocked nucleic acid – an RNA modification with broad potential Bioorg. Med. Chem. 9 (2011) 3591-3597.

**40** A. Novak, S. Hsu, C. Leung-Hagesteijn, G. Radeva, J. Papkoff, R. Montesano, C. Roskelley, R. Grosschedl, S. Dedhar, Cell adhesion and the integrin-linked kinase regulate the LEF-1 and β-catenin signaling pathways, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 4374–4379.

**41** S. Cogoi, V. Rapozzi, S. Cauci, LE. Xodo, Critical role of hnRNP A1 in activating KRAS transcription in pancreatic cancer cells: A molecular mechanism involving G4 DNA, Biochim. Biophys. Acta 1861 (2017) 1389-1398.

**42** PC. Chu, MC. Yang, SK. Kulp, SB. Salunke, LE. Himmel, CS. Fang, AM. Jadhav, YS. Shan, CT. Lee, MD Lai, LA. Shirley, T. Bekaii-Saab, CS. Chen, Regulation of oncogenic KRAS signaling via a novel KRAS-integrin-linked kinase-hnRNPA1 regulatory loop in human pancreatic cancer cells, Oncogene 35 (2016) 3897–3908.

**43** GT. Noble, JF. Stefanick, JD. Asheley, T. Kiziltepe, B. Bilgicer, Ligand-targeted liposome design: challenges and fundamental considerations, Trends Biotechnol 32 (2014) 32-45.

**44** M. Green, PM. Loewenstein, Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein, Cell 55 (1988) 1179-1188.

**45** AD. Frankel, CO. Pabo, Cellular uptake of the tat protein from human immunodeficiency virus, Cell 55 (1988) 1189-1193.

**46** S. Cogoi, U. Jakobsen, E. B. Pedersen, S. Vogel , L. E. Xodo, Lipid-modified G4-

decoy oligonucleotide anchored to nanoparticles: delivery and bioactivity in pancreatic cancer cells, Sci Rep. 6 (2016) 38468.

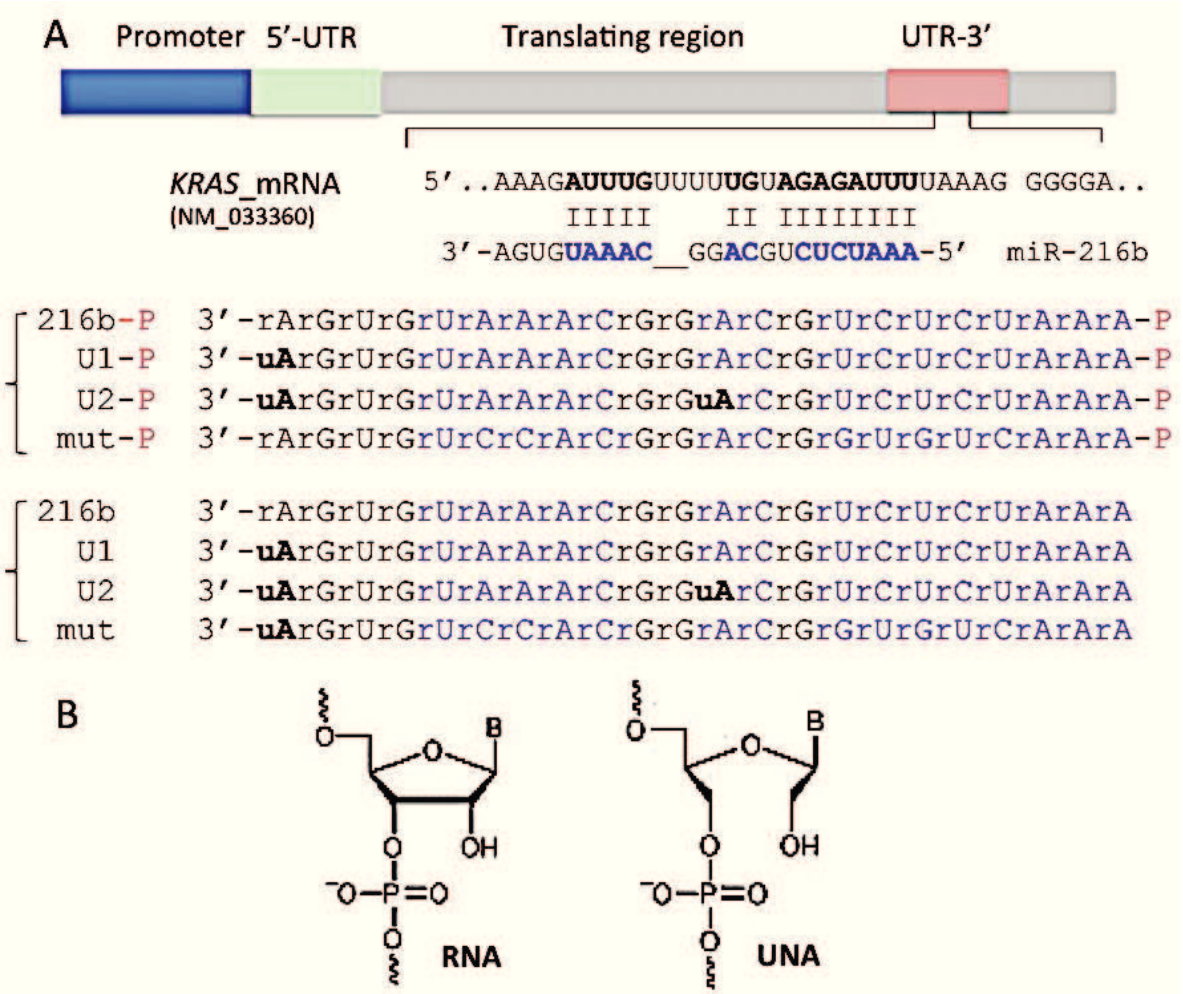


**Figure 1.** (**A**) Sequence of the *KRAS* 3'-UTR recognized by miR-216b (called 216b). We designed single-stranded UNA-modified 216b mimics with and without a phosphate at the 5' end: wild-type 216b and 216b-P; 216b with one UNA (U1 and U1-P) or two UNA (U2 and U2-P) at the 3' end. Mutated miRNAs with UNA (mut) or with a 5' phosphate (mut-P) were used as a control (Supplementary Data, S2); (**B**). The structures of RNA and UNA, lacking the covalent bond between C2' and C3' of the ribose.

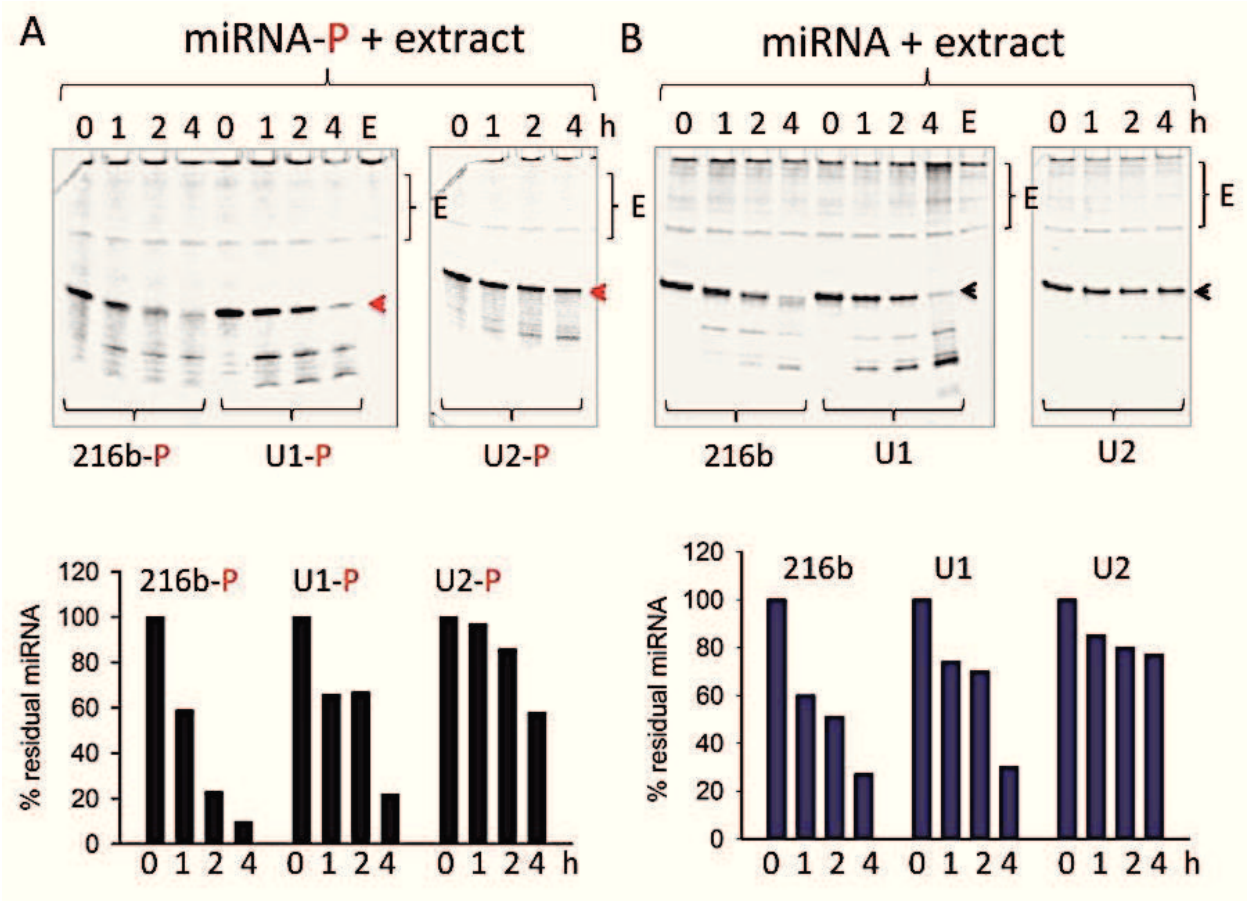


**Figure 2. (A,B)** Denaturing (7 M urea) polyacrylamide gel showing the integrity of wild-type and UNA-modified mimics with and without a 5' phosphate, after incubation for 1, 2 and 4 h with 2 μg Panc-1 cellular extract (E) at 37 °C. The mobility of E is shown in panels 1 and 3 from left. The arrow shows the undegraded miRNA. The % of undegraded miRNA at the various times is shown in the histograms.

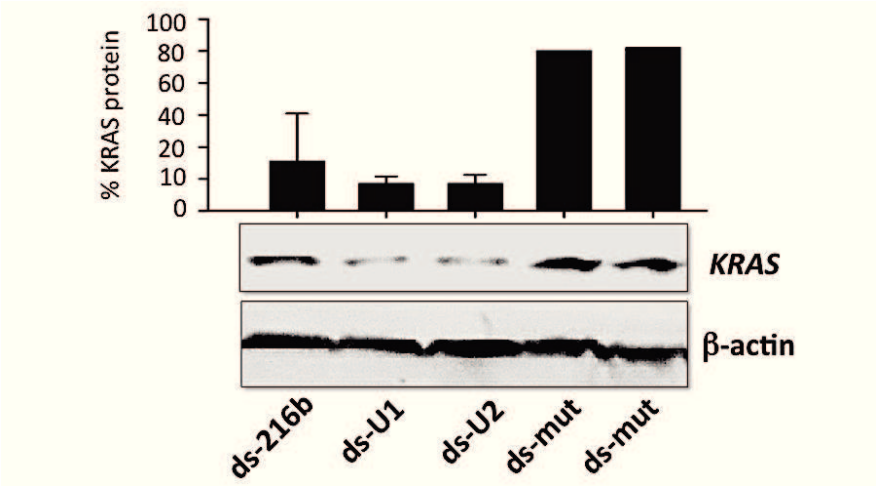



**Figure 3.** Effect of double-stranded UNA-modified mimics (ds-216b, ds-U1, ds-U2, ds-mut) on *KRAS* expression in Panc-1 cells. 216b, U1, U2 and mut were

hybridized to the complementary strand. The Western blot with Panc-1 cells was performed 48 h after the second transfection. The histogram shows the % KRAS protein determined as T/C x 100 where T= (KRAS/actin) in cells treated with the 216b-mimics, C= (KRAS/actin ) in mut-treated cells. Values are the average of two independent experiments.

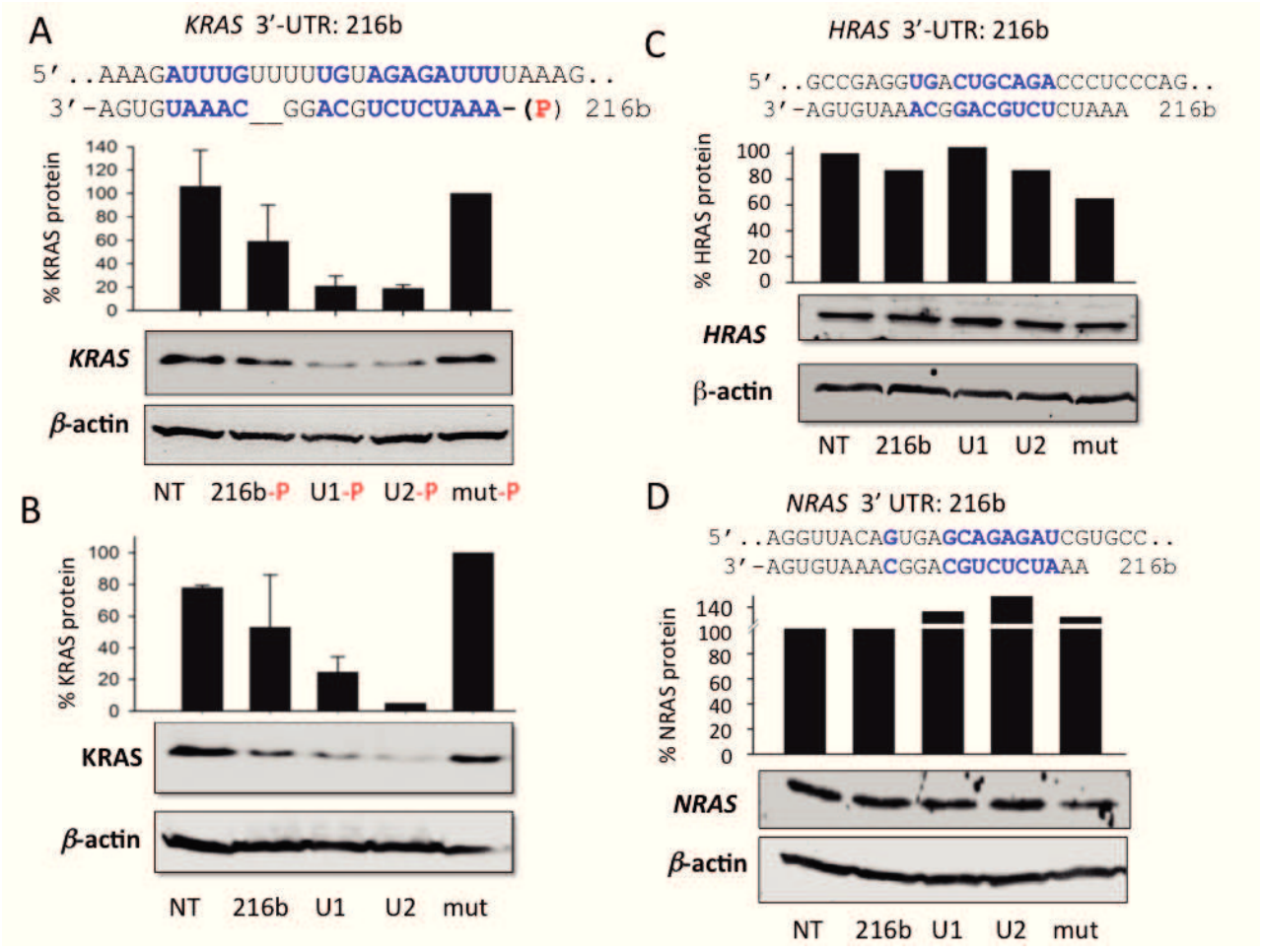



**Figure 4**. Effect of single-stranded miRNA mimics with and without a phosphate at the 5' end on the expression of *KRAS* in Panc-1 cells. The cells, 24 h after seeding, were treated with the mimics, using *Interferin*. A second treatment was carried out 24 h after the first treatment. A Western blot was carried out 48 h after the second treatment. (**A**) Base-pairing between *KRAS* 3'-UTR mRNA and miRNA 216b. The Western blot shows the impact of the designed UNA-modified mimics with a phosphate at the 5' end on protein KRAS. The % KRAS protein in miRNA-treated cells compared to mut-treated cells is reported. The results are the average of three independent experiments; (**B**) as in A but with single-stranded mimics without a phosphate at the 5' end. Results from two

independent experiments; (**C**) and (**D**) Matching between *HRAS* and *NRAS* 3'-UTR mRNA and 216b. The Western blots show the levels of HRAS and NRAS proteins in Panc-1 cells treated with the designed single-stranded miRNAs.

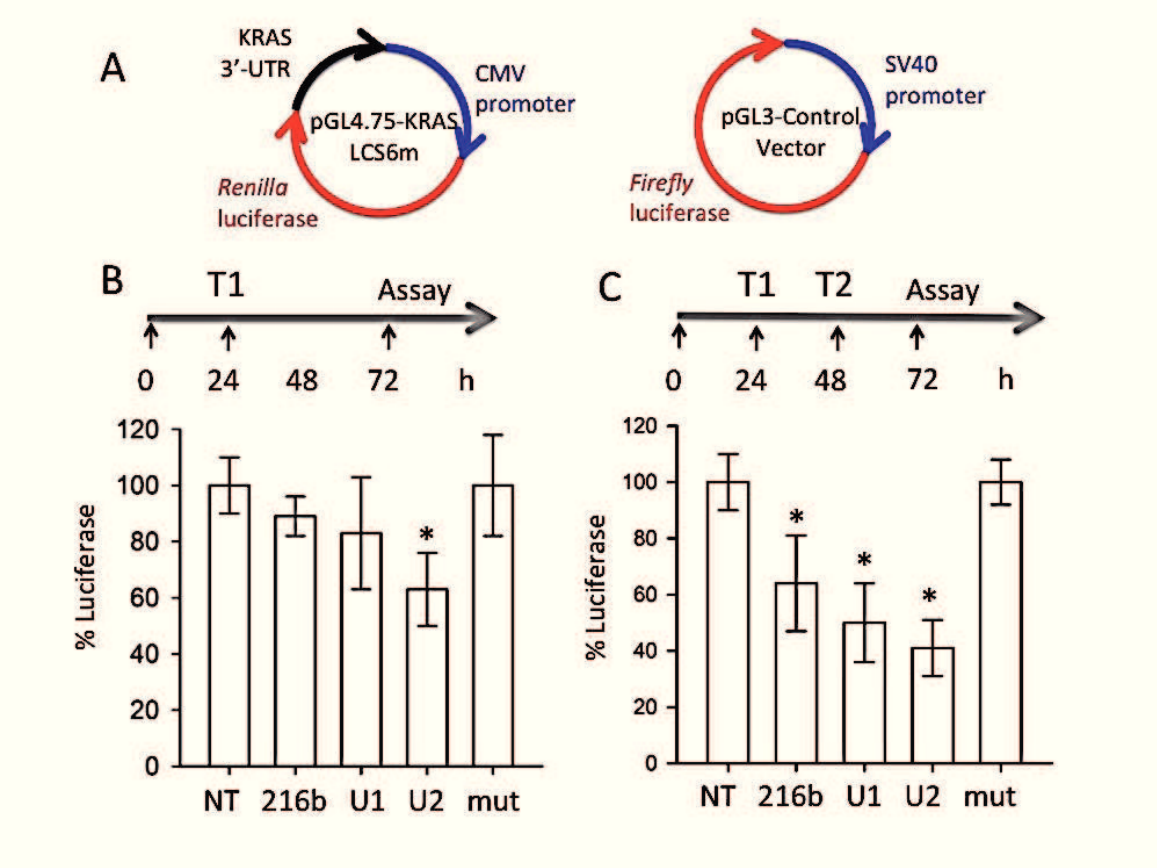



**Figure 5.** (**A**) Plasmids used for the luciferase experiments; (**B**) Outline of the luciferase experiments. T1: plasmids transfection and first miRNA treatment; T2: second miRNA treatment; (**C**) Histograms show the levels of *Renilla* luciferase in Panc-1 cells treated with miR-216b-mut (mut); miR-216b (216b); miR-216b-U1 (U1) and miR-216b-U2 (U2). The ordinate of the histogram reports *Renilla*/*Firefly* in Panc-1 cells treated with miR-216b-mut and with miR-216b or UNA analogue. T/M

x 100 = % *Renilla* luciferase in miRNA treated cells (T) compared to mut-treated cells (M). *=P< 0.05.

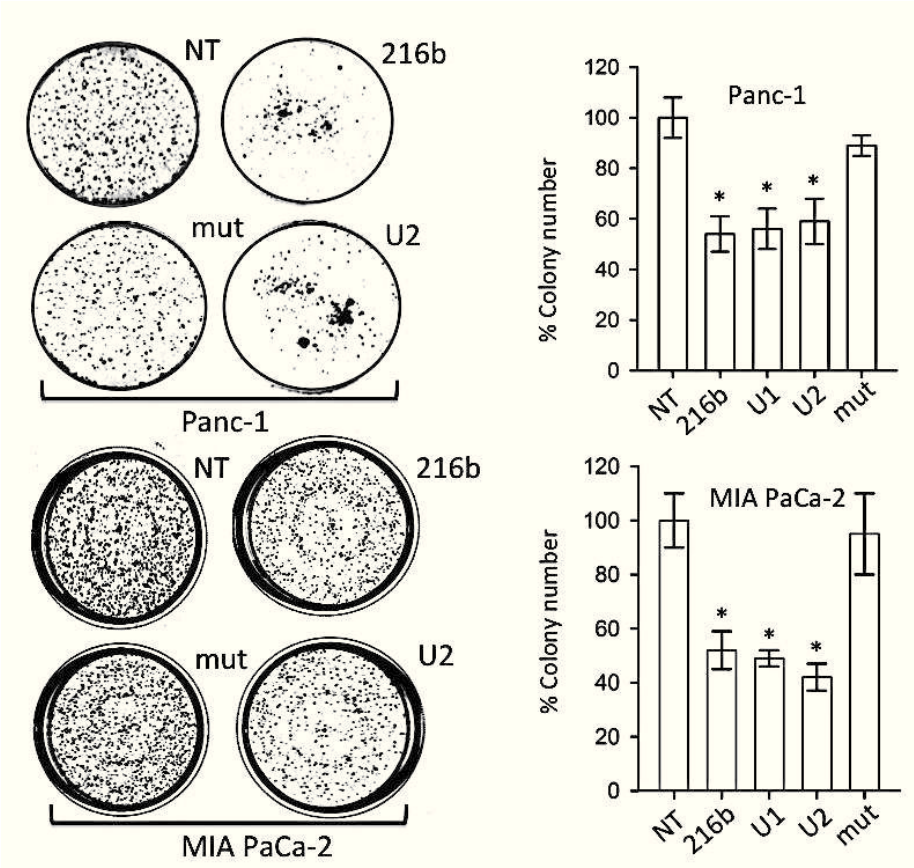



**Figure 6.** Clonogenic assays. MiR-216b (216b) and UNA-modified analogues miR-miR-216b-U1 (U1) and miR-216b-U2 (U2) decrease the % of colonies in Panc-1 and Mia PaCa-2 pancreatic cancer cells. NT= untreated cells. Histogram shows the % colonies in miRNA-treated cells compared to NT. Error bar is the average of three independent experiments. *=P < 0.05. As for Panc-1 cells the designed miRNAs reduced the number of colonies by at least 50 %.

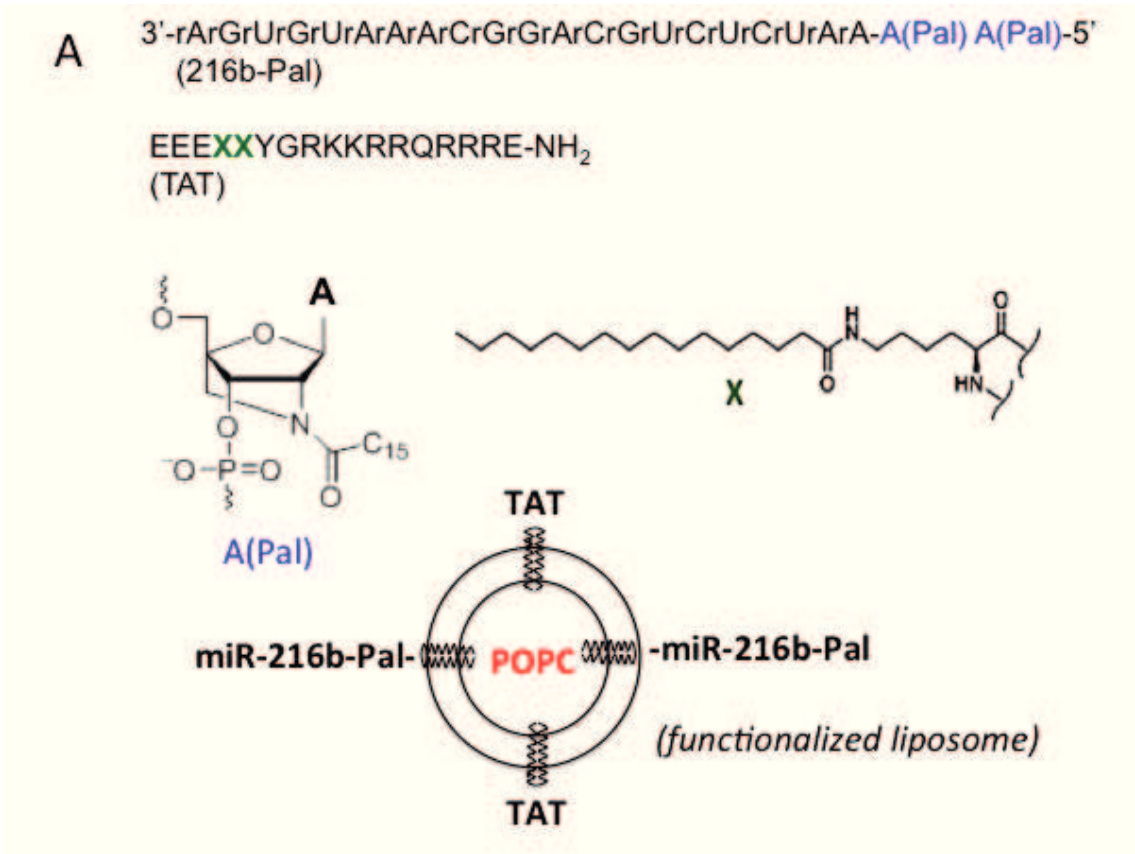
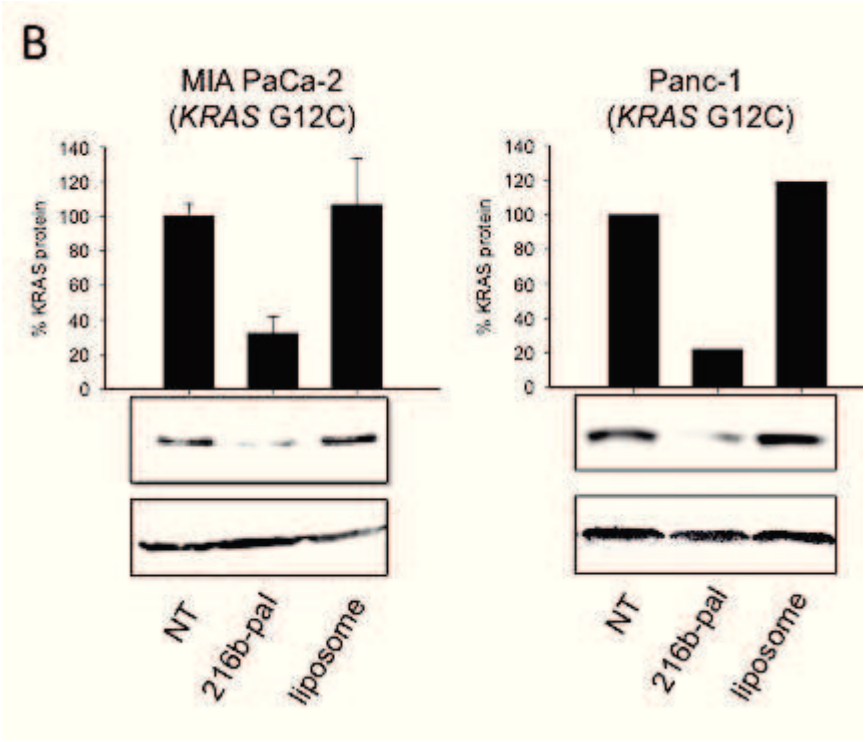



**Figure 7.** (**A**) Sequence of miR-216b with two palmitoyl chains and of the TAT peptide with two **X** insertions containing two saturated palmityl chains in order to anchor the peptide to



the liposomes. MiR-216b is chemically modified with two A(Pal) residues, each containing a saturated palmityl chains in order to anchor the oligonucleotides to the liposomes. The POPC liposomes spontaneously self-assembles into spherically closed bilayer membrane on the surface of which TAT and 216b-Pal are attached through their lipid modifications.

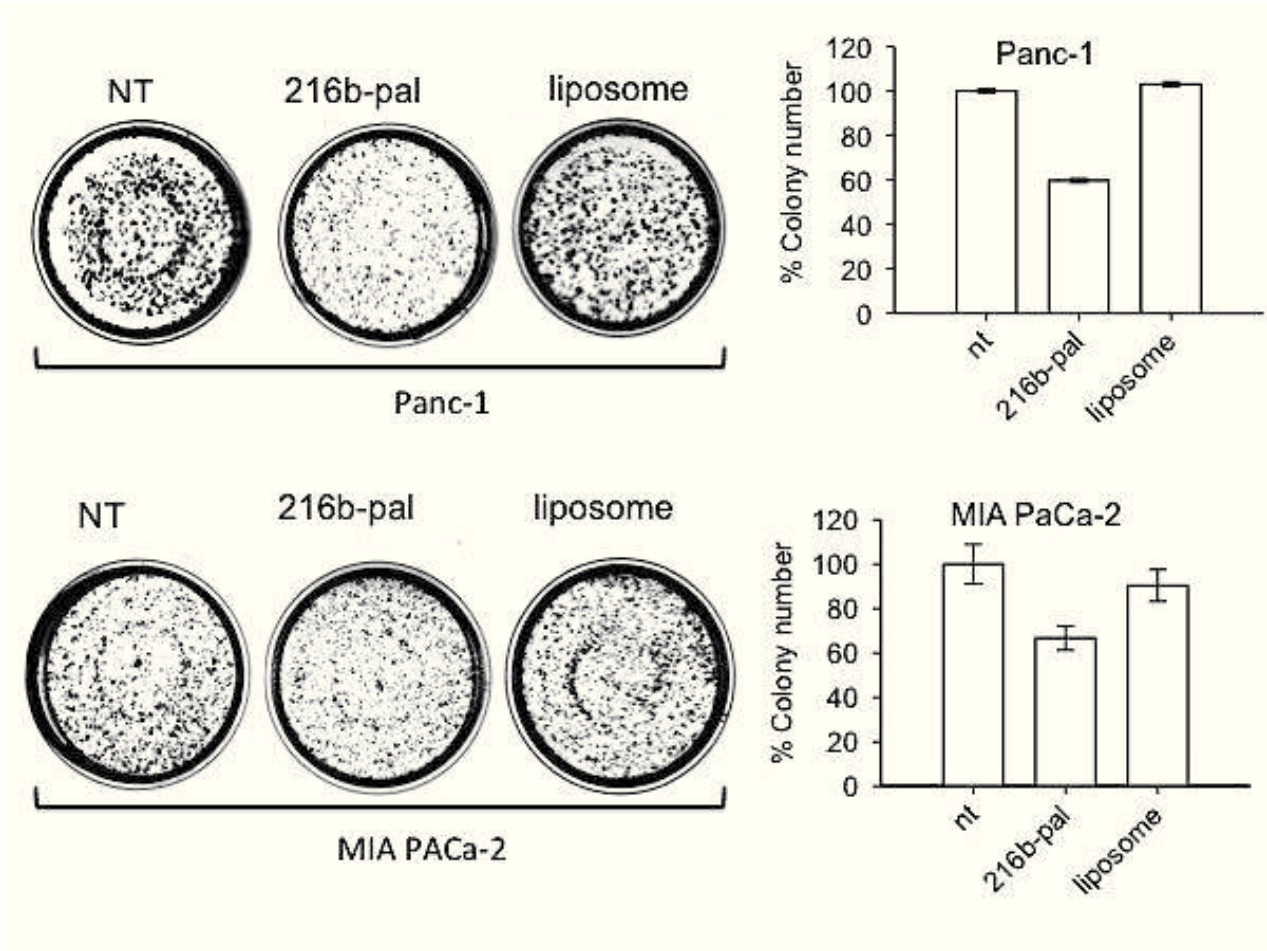



**Figure 8.** Clonogenic assays. POPC liposome functionalized with 216b-pal decrease the % of colonies in Panc-1 and Mia PaCa-2 pancreatic cancer cells. NT= untreated cells, 216-pal= cells treated with POPC liposomes functionalized with 216b-pal and TAT; Liposome= cells treated with POPC non-functionalized with the effector molecules. Histogram shows the % colonies in miRNA-treated, liposome-treated and non treated (NT) cells. Error bar is the average of three independent experiments.

SUPPLEMENTARY DATA

# Potent anti Kirsten ras (*KRAS*) single-stranded miRNA oligonucleotide mimics as therapeutic agents.

Giulia Miglietta,[a] Annalisa Ferino,[a] Susanna Cogoi,[a] Raffaella Picco,[a] Stefan Vogel,[b] Jesper Wengel [b] and Luigi E. Xodo [a] *

(a) Department of Medicine, Laboratory of Biochemistry, University of Udine, Italy; **luigi.xodo@uniud.it**;

(b) Nucleic Acids Centre, University of Southern Denmark, DK 5230, Odense, Denmark

**Supplementary Data, S1.** miRNA and FC values calculated for three dataset (GSE43796, E-MTAB-753 and E-GEOD-41372).

## GSE43796

| miRNA | FC |
|---|---|
| hsa-miR-216a | 36.83 |
| hsa-miR-216a | 26.34 |
| hsa-miR-216a | 15.82 |
| hsa-miR-216a | 3.25 |
| **hsa-miR-216b** | **27.95** |
| **hsa-miR-216b** | **13.03** |
| **hsa-miR-216b** | **8.83** |
| hsa-miR-217 | 21.26 |
| hsa-miR-217 | 13.73 |
| hsa-miR-217 | 8.61 |
| hsa-miR-217 | 2.25 |
| hsa-miR-30c | 2.09 |
| hsa-miR-30c | 1.76 |
| hsa-miR-30e | 2.05 |
| hsa-miR-30e | 1.59 |
| hsa-miR-30b | 1.86 |
| hsa-miR-30b | 1.56 |
| hsa-miR-30b | 1.44 |
| hsa-miR-30a | 1.85 |
| hsa-miR-181a | -2.64 |
| hsa-miR-181b | -2.63 |
| hsa-miR-181b | -2.35 |
| hsa-miR-200b | -2.49 |
| hsa-miR-143 | -1.71 |
| hsa-miR-143 | -1.37 |
| hsa-miR-1 | -1.40 |
| hsa-miR-1 | -1.06 |
| hsa-miR-155 | -1.69 |
| hsa-miR-155 | -1.47 |

## E-MTAB-753

| miRNA | FC |
|---|---|
| hsa-miR-217 | 2.95 |
| **hsa-miR-216b** | **2.27** |
| hsa-miR-30a | 1.77 |
| hsa-miR-30b | 1.76 |
| hsa-miR-30d | 1.59 |
| hsa-miR-143 | -6.33 |
| hsa-miR-1 | -1.24 |
| hsa-miR-199b-3p | -4.79 |
| hsa-miR-199a-3p | -4.61 |
| hsa-miR-181a | -4.37 |
| hsa-miR-155 | -3.90 |

## E-GEOD-41372

| miRNA | FC |
|---|---|
| hsa-miR-217 | 14.67 |
| hsa-miR-216a | 8.96 |
| hsa-miR-30d | 1.67 |
| hsa-miR-30a | 1.63 |
| hsa-miR-30e | 1.60 |

**Fold Change = FC = $Log_2$ (F ctrl /F tumor)**

**Table S2.** Oligonucleotides used in this study.

| | |
|---|---|
| 216b | 5'-rArArArUrCrUrCrUrGrCrArGrGrCrArArArUrGrUrGrA |
| U1 | 5'-rArArArUrCrUrCrUrGrCrArGrGrCrArArArUrGrUrGuA |
| U2 | 5'-rArArArUrCrUrCrUrGrCuArGrGrCrArArArUrGrUrGuA |
| mut | 5'-rArArArCrUrGrUrGrGrCrArGrGrCrArCrCrUrGrUrGuA |
| 216b-P | 5'P-rArArArUrCrUrCrUrGrCrArGrGrCrArArArUrGrUrGrA |
| U1-P | 5'P-rArArArUrCrUrCrUrGrCrArGrGrCrArArArUrGrUrGuA |
| U2-P | 5'P-rArArArUrCrUrCrUrGrCuArGrGrCrArArArUrGrUrGuA |
| mut-P | 5'P-rArArArCrUrGrUrGrGrCrArGrGrCrArCrCrUrGrUrGrA |
| comlementary-216b | 5'-rCrArCrArUrUrUrGrCrCrUrGrCrArGArGrArUrUrUrU |
| complementary-mut: | 5'-rCrArCrArGrGrUrGrCrCrUrGrCrCrArCrArGrUrUrUrU |

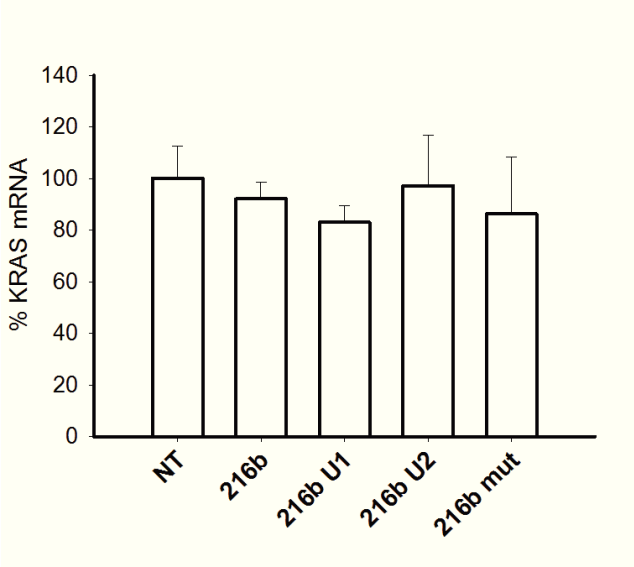
uA= unlocked nucleic acid modification;

**Table S3.** Molecular masses of UNA-modified RNAs by MALDI MS.

| miRNA sequence* | Name | Calculated | Found |
|---|---|---|---|
| 5'-rArArArUrCrUrCrUrGrCrArGrGrCrArArArUrGrUrGuA | U1 | 7051,4 | 7051,3 |
| 5'-rArArArUrCrUrCrUrGrCuArGrGrCrArArArUrGrUrGuA | U2 | 7053,4 | 7053,4 |
| 5'-rArArArCrUrGrUrGrGrCrArGrGrCrArCrCrUrGrUrGuA | mut | 7082,4 | 7082,7 |
| 5'P-rArArArUrCrUrCrUrGrCrArGrGrCrArArArUrGrUrGuA | U1-P | 7131,3 | 7131,3 |
| 5'P-rArArArUrCrUrCrUrGrCuArGrGrCrArArArUrGrUrGuA | U2-P | 7133,3 / | 7133,6 |

* uA= unlocked nucleic acid

**Supplementary Data, S4.** Determination of the level of KRAS mRNA in Panc-1 cells by quantitative real-time PCR. The assay has been performed after 24h of the miR-216b and UNA-modified analogs transfection.



# 4. CONCLUSION

The main aim of my PhD work has been to investigate the impact of unusual DNA structures formed in the ras promoter and their involvement in biological processes such as transcription and translation. Moreover I focused on the development of innovative ras-specific anticancer strategies. As the expression of the ras genes is crucial for cancer initiation and maintenance, they are attractive therapeutic targets. The identification and characterization of ras non-canonical DNA structures is a pre-requisite to develop alternative anti-ras strategies.

First, I focused on the *HRAS* promoter. I have investigated the C-rich strands of two C-rich regions near TSS, to understand their ability to fold into a DNA unusual structure named *i*-motif (*i*M). My work demonstrated that the *HRAS* C-rich strands fold into the *i*M under slightly acidic conditions and under crowding conditions. Such *i*M conformation provides a chemical scaffold that is recognized by hnRNP A1 nuclear factor. I found that this protein binds tightly to the *i*M and, by means of FRET experiments, I also found that it unfolds the *i*M *in vitro*. We also address the question of elucidating the possible role of hnRNP A1 in *HRAS* transcription. Based on my work and data previously obtained in my laboratory, I have proposed a dynamic molecular G-quadruplex/*i*M switch for the control of *HRAS* expression. Secondly, I have proposed and discussed a decoy strategy for repressing *HRAS* transcription in bladder cancer cells. Previous studies showed that *HRAS* expression is regulated by two neighbouring G-quadruplexes (*hras-1* and *hras-2*) near TSS [117]. Their biological function is to promote the recruitment of proteins such as MAZ and hnRNP A1 to the promoter G4 motifs. These proteins should unfold the G-quadruplexes and thus activate transcription. We designed a decoy strategy to inhibit HRAS in bladder cancer cells. The oligonucleotides carried chemical modifications (anthraquinone insertions and LNA nucleic acid modification) to increase their nuclease resistance. I designed decoy oligonucleotides mimicking either the *i*M or the G4 structures of the *HRAS* promoter. The decoy oligonucleotides effectively decreased the recruitment of MAZ to the *HRAS* promoter and inhibited transcription.

The last part of my PhD program focused on KRAS, specifically on the 5'- and 3'-untranslated regions (UTRs). The KRAS 5'-UTR is characterized by a high GC content (77 %). By using bioinformatic tools we evaluated and selected sequences with a high propensity to form RNA G-quadruplex structures (G4s). By using different techniques I have demonstrated the presence of non-overlapping RG4s in the first 80 nt of the *KRAS* 5'-UTR (Circular Dicroism, RNase footprinting, Electrophoretic Mobility Shift Assay, UV melting curves and Thioflavin T assay). In addition, I have developed an efficient anti-*KRAS* strategy by using small molecules (anthraquinone derivatives) that easily penetrate the cells and bind to the *KRAS* RG4s. I have developed a pull-down assay to demonstrate that the small

molecules effectively recognize and bind to the *KRAS* transcripts under low abundance cellular condition. Anthraquinone derivatives, upon binding to the KRAS RG4s, showed to be able to downregulate the *KRAS* expression. As pancreatic cancer cells are addicted to *KRAS*, the decrease of KRAS protein promoted apoptosis and a strong antiproliferative effect in PDAC cells. Simultaneously we also proposed an anti-*KRAS* strategy targeting the 3'-UTR of the *KRAS* transcript. Starting from the observation that in pancreatic tumours miR-216b, of which *KRAS* is a target, is abnormally downregulated, we proposed an alternative strategy based on restoring the level of this miRNA to inhibit *KRAS* translation. We used single stranded RNA mimics of miR-216b which were engineered with unlocked nucleic acid (UNA) modifications to improved their nuclease resistance *in vivo*. The transfections of pancreatic cancer cells with the designed miRNAs mimics resulted in a significant reduction of KRAS protein level. The designed miRNAs decreased the number of colonies in two different pancreatic cell lines. We also tested a new delivery strategy based on the use of palmitoyl-oleyl-phosphatidylcholine (POPC) liposomes functionalized with lipid-modified miR-216b and lipid-modified cell penetrating TAT peptide.

**BIBLIOGRAPHY**


[1] J. J. Harvey, "An unidentified virus which causes the rapid production of tumours in mice," *Nature*, vol. 204, no. 4963, pp. 1104–1105, 1964.

[2] W. H. Kirsten and L. A. Mayer, "Morphologic Responses to a Murine Erythroblastosis Virus 2," *J. Natl. Cancer Inst.*, vol. 39, no. 2, pp. 311–335, 1967.

[3] E. M. Scolnick and W. P. Parks, "Harvey sarcoma virus: a second murine type C sarcoma virus with rat genetic information.," *J. Virol.*, vol. 13, no. 6, pp. 1211–9, 1974.

[4] E. M. Scolnick, E. Rands, D. Williams, and W. P. Parks, "Studies on the nucleic acid sequences of Kirsten sarcoma virus: a model for formation of a mammalian RNA-containing sarcoma virus.," *J. Virol.*, vol. 12, no. 3, pp. 458–63, 1973.

[5] T. Y. Shih, M. O. Weeks, H. A. Young, and E. M. Scholnick, "Identification of a sarcoma virus-coded phosphoprotein in nonproducer cells transformed by Kirsten or Harvey murine sarcoma virus," *Virology*, vol. 96, no. 1, p. 64—79, 1979.

[6] E. M. Scolnick, A. G. Papageorge, and T. Y. Shih, "Guanine nucleotide-binding activity as an assay for src protein of rat-derived murine sarcoma viruses," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 76, no. 10, p. 5355—5359, 1979.

[7] M. C. Willingham, I. Pastan, T. Y. Shih, and E. M. Scolnick, "Localization of the src gene product of the Harvey strain of MSV to plasma membrane of transformed cells by electron microscopic immunocytochemistry," *Cell*, vol. 19, no. 4, p. 1005—1014, Apr. 1980.

[8] M. Janakiraman, E. Vakiani, Z. Zeng, C. A. Pratilas, B. S. Taylor, D. Chitale, E. Halilovic, M. Wilson, K. Huberman, J. C. Ricarte Filho, Y. Persaud, D. A. Levine, J. A. Fagin, S. C. Jhanwar, J. M. Mariadason, A. Lash, M. Ladanyi, L. B. Saltz, A. Heguy, P. B. Paty, and D. B. Solit, "Genomic and biological characterization of exon 4 KRAS mutations in human cancer," *Cancer Res.*, vol. 70, no. 14, pp. 5901–5911, 2010.

[9] C. Shih, B.-Z. Shilo, M. P. Goldfarb, A. Dannenberg, and R. A. Weinberg, "Passage of phenotypes of chemically transformed cells via transfection of DNA and chromatin (chemical carcinogenesis/transformation alleles/Southern blotting)," *Cell Biol.*, vol. 76, no. 11, pp. 5714–5718, 1979.

[10] T. G. Krontiris and G. M. Cooper, "Transforming activity of human tumor DNAs.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 78, no. 2, pp. 1181–4, 1981.

[11] M. J. Murray, B. Z. Shilo, C. Shih, D. Cowing, H. W. Hsu, and R. a Weinberg, "Three different human tumor cell lines contain different oncogenes.," *Cell*, vol. 25, no. 2, pp. 355–361, 1981.

[12] C. Shih, L. C. Padhy, M. Murray, and R. A. Weinberg, "Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts," *Nature*, vol. 290, no. 5803, p. 261—264, 1981.

[13] M. Perucho, M. Goldfarb, K. Shimizu, and M. Wigler, "Human-Tumor-Derived Common and Different Cell Lines Contain Transforming Genes," vol. 27, no. December, pp. 467–476, 1981.

[14] C. J. Marshall, A. Hall, and R. A. Weiss, "A transforming gene present in human sarcoma cell lines," *Nature*, vol. 299, no. 5879, p. 171—173, 1982.

[15] E. Santos, S. R. Tronick, S. A. Aaronson, S. Pulciani, and M. Barbacid, "T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of BALB- and Harvey-MSV transforming genes.," *Nature*, vol. 298, no. 5872, pp. 343–347, 1982.

[16] L. F. Parada, C. J. Tabin, C. Shih, and R. A. Weinberg, "Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene," *Nature*, vol. 297, no. 5866, pp. 474–478, Jun. 1982.

[17] E. Taparowsky, Y. Suard, O. Fasano, K. Shimizu, M. Goldfarb, and M. Wigler, "Activation of the T24 bladder carcinoma transforming gene is linked to a single amino acid change," *Nature*, vol. 300, no. 5894, p. 762—765, 1982.

[18] C. J. Tabin, S. M. Bradley, C. I. Bargmann, R. A. Weinberg, A. G. Papageorge, E. M. Scolnick, R. Dhar, D. R. Lowy, and E. H. Chang, "Mechanism of activation of a human oncogene," *Nature*, vol. 300, no. 5888, pp. 143–149, 1982.

[19] K. Shimizu, M. Goldfarb, Y. Suard, M. Perucho, Y. Li, T. Kamata, J. Feramisco, E. Stavnezer, J. Fogh, and M. H. Wigler, "Three human transforming genes are related to the viral ras oncogenes.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 80, no. 8, pp. 2112–6, 1983.

[20] A. Hall, C. J. Marshall, N. K. Spurr, and R. A. Weiss, "Identification of transforming gene in two human sarcoma cell lines as a new member of the ras gene family located on chromosome 1," *Nature*, vol. 303, no. 5916, p. 396—400, 1983.

[21] P. a Boriack-Sjodin, S. M. Margarit, D. Bar-Sagi, and J. Kuriyan, "The structural basis of the activation of Ras by Sos.," *Nature*, vol. 394, no. 6691, pp. 337–343, 1998.

[22] J. F. Hancock, "Ras proteins: different signals from different locations," *Nat. Rev. Mol. Cell Biol.*, vol. 4, no. 5, pp. 373–385, 2003.

[23] M. E. Furth, T. H. Aldrich, and C. Cordon-Cardo, "Expression of ras proto-oncogene proteins in normal human tissues," *Oncogene*, vol. 1, no. 1, p. 47—58, 1987.

[24] J. Leon, I. Guerrero, and A. Pellicer, "Differential expression of the ras gene family in mice," *Mol. Cell. Biol.*, vol. 7, no. 4, p. 1535—1540, Apr. 1987.

[25] G. Fiorucci and A. Hall, "All three human ras genes are expressed in a wide range of tissues," *Biochim. Biophys. Acta*, vol. 950, no. 1, p. 81—83, 1988.

[26] J. Leon, I. Guerrero, and a Pellicer, "Differential expression of the ras gene family in mice.," *Mol. Cell. Biol.*, vol. 7, no. 4, pp. 1535–40, 1987.

[27] L. Johnson, D. Greenbaum, K. Cichowski, K. Mercer, E. Murphy, E. Schmitt, R. T. Bronson, H. Umanoff, W. Edelmann, R. Kucherlapati, and T. Jacks, "K-ras is an essential gene in the mouse with partial functional overlap with N-ras," *Genes Dev.*, vol. 11, no. 19, pp. 2468–2481, 1997.

[28] D. R. Lowy and B. M. Willumsen, "Function and Regulation of Ras," *Annu. Rev. Biochem.*, vol. 62, no. 1, pp. 851–891, Jun. 1993.

[29] B. M. Willumsen, A. Christensen, N. L. Hubbert, A. G. Papageorge, and D. R. Lowy, "The p21 ras C-terminus is required for transformation and membrane association," *Nature*, vol. 310, no. 5978, pp. 583–586, Aug. 1984.

[30] J. F. Hancock, H. Paterson, and C. J. Marshall, "A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize p21ras to the plasma membrane," *Cell*, vol.

63, no. 1, pp. 133–139, 1990.

[31] M. M. McKay and D. K. Morrison, “Integrating signals from RTKs to ERK/MAPK,” *Oncogene*, vol. 26, no. 22, pp. 3113–3121, 2007.

[32] F. D. Tsai, M. S. Lopes, M. Zhou, H. Court, O. Ponce, J. J. Fiordalisi, J. J. Gierut, A. D. Cox, K. M. Haigis, and M. R. Philips, “K-Ras4A splice variant is widely expressed in cancer and uses a hybrid membrane-targeting motif,” *Proc. Natl. Acad. Sci.*, vol. 112, no. 3, pp. 779–784, 2015.

[33] I. A. Prior and J. F. Hancock, “Ras trafficking, localization and compartmentalized signalling,” *Semin. Cell Dev. Biol.*, vol. 23, no. 2, pp. 145–153, 2012.

[34] S. L. Moores, M. D. Schaber, S. D. Mosser, E. Rands, M. B. O’Hara, V. M. Garsky, M. S. Marshall, D. L. Pompliano, and J. B. Gibbs, “Sequence dependence of protein isoprenylation,” *J. Biol. Chem.*, vol. 266, no. 22, pp. 14603–14610, 1991.

[35] D. B. Whyte, P. Kirschmeier, T. N. Hockenberry, I. Nunez-Oliva, L. James, J. J. Catino, W. R. Bishop, and J. K. Pai, “K- and N-Ras are geranylgeranylated in cells treated with farnesyl protein transferase inhibitors,” *J. Biol. Chem.*, vol. 272, no. 22, pp. 14459–14464, 1997.

[36] L. Johnson, D. Greenbaum, K. Cichowski, K. Mercer, E. Murphy, E. Schmitt, R. T. Bronson, H. Umanoff, W. Edelmann, R. Kucherlapati, and T. Jacks, “K- ras is an essential gene in the mouse with partial functional overlap with N- ras,” *Genes Dev.*, vol. 11, no. Bos 1989, pp. 2468–2481, 1997.

[37] K. Koera, K. Nakamura, K. Nakao, J. Miyoshi, K. Toyoshima, T. Hatta, H. Otani, A. Aiba, and M. Katsuki, “K-Ras is essential for the development of the mouse embryo,” *Oncogene*, vol. 15, no. 10, pp. 1151–1159, 1997.

[38] N. Potenza, C. Vecchione, A. Notte, A. De Rienzo, A. Rosica, L. Bauer, A. Affuso, M. De Felice, T. Russo, R. Poulet, G. Cifelli, G. De Vita, G. Lembo, and R. Di Lauro, “Replacement of K-Ras with H-Ras supports normal embryonic development despite inducing cardiovascular pathology in adult mice.,” *EMBO Rep.*, vol. 6, no. 5, pp. 432–7, 2005.

[39] H. Umanoff, W. Edelmann, A. Pellicer, and R. Kucherlapati, “The murine N-ras gene is not essential for growth and development.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 92, no. 5, pp. 1709–1713, 1995.

[40] P. P. Fotiadou, C. Takahashi, H. N. Rajabi, and M. E. Ewen, “Wild-type NRas and KRas perform distinct functions during transformation.,” *Mol. Cell. Biol.*, vol. 27, no. 19, pp. 6742–6755, 2007.

[41] R. Khosravi-Far and C. J. Der, “The Ras signal transduction pathway,” *Cancer Metastasis Rev.*, vol. 13, no. 1, pp. 67–89, 1994.

[42] Lefort, “Ras activation by insulin and epidermal growth factor through enhanced of guanine nucleotides on p21ras,” *Mol. Cell. Biol.*, vol. 13, no. 1, pp. 155–162, 1993.

[43] S. Traverse, N. Gomez, H. Paterson, C. Marshall, and P. Cohen, “Sustained activation of the mitogen-activated protein (MAP) kinase cascade may be required for differentiation of PC12 cells. Comparison of the effects of nerve growth factor and epidermal growth factor.,” *Biochem. J.*, vol. 288 ( Pt 2, pp. 351–5, 1992.

[44] C.-H. Heldin, “Dimerization of cell surface receptors in signal transduction,” *Cell*, vol. 80, no. 2, pp. 213–223, 1995.

[45] C. . Marshall, “Specificity of receptor tyrosine kinase signaling: Transient versus sustained extracellular signal-regulated kinase activation,” *Cell*, vol. 80, no. 2, pp. 179–185, 1995.

[46] K. Rajalingam, R. Schreck, U. R. Rapp, and S. Albert, “Ras oncogenes and their downstream targets,” *Biochim. Biophys. Acta - Mol. Cell Res.*, vol. 1773, no. 8, pp. 1177–1195, 2007.

[47] J. Schlessinger, “Cell Signaling by Receptor Tyrosine Kinases A large group of genes in all eukaryotes encode for,” *October*, vol. 103, no. 2, pp. 211–225, 2000.

[48] P. Gideon, J. John, M. Frech, a Lautwein, R. Clark, J. E. Scheffler, and a Wittinghofer, “Mutational and kinetic analyses of the GTPase-activating protein (GAP)-p21 interaction: the C-terminal domain of GAP is not sufficient for full activity.,” *Mol. Cell. Biol.*, vol. 12, no. 5, pp. 2050–2056, 1992.

[49] M. Malumbres and M. Barbacid, “RAS oncogenes: the first 30 years,” *Nat. Rev. Cancer*, vol. 3, no. 6, pp. 459–465, 2003.

[50] R. Ballester, D. Marchuk, M. Boguski, A. Saulino, R. Letcher, M. Wigler, and F. Collins, “The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins,” *Cell*, vol. 63, no. 4, pp. 851–859, 1990.

[51] G. Xu, P. O’Connell, D. Viskochil, R. Cawthon, M. Robertson, M. Culver, D. Dunn, J. Stevens, R. Gesteland, R. White, and R. Weiss, “The neurofibromatosis type 1 gene encodes a protein related to GAP,” *Cell*, vol. 62, no. 3, pp. 599–608, 1990.

[52] A. Bernards and J. Settleman, “GAP control: Regulating the regulators of small GTPases,” *Trends Cell Biol.*, vol. 14, no. 7, pp. 377–385, 2004.

[53] G. A. Repasky, E. J. Chenette, and C. J. Der, “Renewing the conspiracy theory debate: Does Raf function alone to mediate Ras oncogenesis?,” *Trends Cell Biol.*, vol. 14, no. 11, pp. 639–647, 2004.

[54] H. Chong, H. G. Vikis, and K. L. Guan, “Mechanisms of regulating the Raf kinase family,” *Cell. Signal.*, vol. 15, no. 5, pp. 463–469, 2003.

[55] J. T. Winston, S. R. Coats, Y. Z. Wang, and W. J. Pledger, “Regulation of the cell cycle machinery by oncogenic ras,” *Oncogene*, vol. 12, no. 1, p. 127—134, 1996.

[56] J. Yan, S. Roy, A. Apolloni, A. Lane, and J. F. Hancock, “Ras isoforms vary in their ability to activate Raf-1 and phosphoinositide 3-kinase,” *J. Biol. Chem.*, vol. 273, no. 37, pp. 24052–24056, 1998.

[57] M. Cully, H. You, A. J. Levine, and T. W. Mak, “Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis,” *Nat. Rev. Cancer*, vol. 6, no. 3, pp. 184–192, 2006.

[58] M. P. Scheid and J. R. Woodgett, “PKB/AKT: functional insights from genetic models.,” *Nat. Rev. Mol. Cell Biol.*, vol. 2, no. 10, pp. 760–768, 2001.

[59] M. G. Kazanietz and M.-J. Caloca, “The Rac GTPase in cancer: from old concepts to new paradigms,” *Cancer Res.*, p. canres.1456.2017, 2017.

[60] L. A. Feig, T. Urano, and S. Cantor, “Evidence for a Ras/Ral signaling cascade,” *Trends Biochem. Sci.*, vol. 21, no. 11, pp. 438–441, 1996.

[61] L. A. Quilliam, J. F. Rebhun, and A. F. B. T.-P. in N. A. R. and M. B. Castro, “A growing family

of guanine nucleotide exchange factors is responsible for activation of ras-family GTPases," vol. 71, no. Supplement C, Academic Press, 2002, pp. 391–444.

[62] Y. Chien and M. A. White, "RAL GTPases are linchpin modulators of human tumour-cell proliferation and survival.," *EMBO Rep.*, vol. 4, no. 8, pp. 800–6, 2003.

[63] L. A. Feig, "Ral-GTPases: Approaching their 15 minutes of fame," *Trends Cell Biol.*, vol. 13, no. 8, pp. 419–425, 2003.

[64] J. L. Bos, "ras Oncogenes in Human Cancer : A Review," pp. 4682–4689, 1989.

[65] I. A. Prior, P. D. Lewis, and C. Mattos, "A comprehensive survey of ras mutations in cancer," *Cancer Res.*, vol. 72, no. 10, pp. 2457–2467, 2012.

[66] R. Shepherd, S. A. Forbes, D. Beare, S. Bamford, C. G. Cole, S. Ward, N. Bindal, P. Gunasekaran, M. Jia, C. Y. Kok, K. Leung, A. Menzies, A. P. Butler, J. W. Teague, P. J. Campbell, M. R. Stratton, and P. A. Futreal, "Data mining using the catalogue of somatic mutations in cancer biomart," *Database*, vol. 2011, no. October, pp. 2–7, 2011.

[67] K. Scheffzek, "The Ras-RasGAP Complex: Structural Basis for GTPase Activation and Its Loss in Oncogenic Ras Mutants," *Science (80-. ).*, vol. 277, no. 5324, pp. 333–338, 1997.

[68] P. H. Seeburg, W. W. Colby, D. J. Capon, D. V Goeddel, and A. D. Levinson, "Biological properties of human c-Ha-ras1 genes mutated at codon 12," *Nature*, vol. 312, no. 5989, p. 71—75, 1984.

[69] S. M. Franken, A. J. Scheidig, U. Krengel, H. Rensland, A. Lautwein, M. Geyer, K. Scheffzek, R. S. Goody, H. R. Kalbitzer, E. F. Pai, and A. Wittinghofer, "Three-Dimensional Structures and Properties of a Transforming and a Nontransforming Glycine-12 Mutant of p21H-ras," *Biochemistry*, vol. 32, no. 33, pp. 8411–8420, 1993.

[70] W. W. Colby, J. S. Hayflick, S. G. Clark, and a D. Levinson, "Biochemical characterization of polypeptides encoded by mutated human Ha-ras1 genes.," *Mol. Cell. Biol.*, vol. 6, no. 2, pp. 730–734, 1986.

[71] M. Castillo-Martin, J. Domingo-Domenech, O. Karni-Schmidt, T. Matos, and C. Cordon-Cardo, "Molecular pathways of urothelial development and bladder tumorigenesis," *Urol. Oncol. Semin. Orig. Investig.*, vol. 28, no. 4, pp. 401–408, 2010.

[72] J. I. E. N. Eble, G. Sauter and I. A. Sesterhenn, "Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs," *Res. Policy*, vol. 34, no. 2, pp. 249–250, 2005.

[73] F. M. Torti, B. L. Lum, D. Aston, N. MacKenzie, M. Faysel, L. D. Shortliffe, and F. Freiha, "Superficial bladder cancer: the primacy of grade in the development of invasive disease.," *J. Clin. Oncol.*, vol. 5, no. 1, pp. 125–130, Jan. 1987.

[74] E. Santos, S. R. Tronick, S. A. Aaronson, S. Pulciani, and M. Barbacid, "T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of BALB- and Harvey-MSV transforming genes," *Nature*, vol. 298, no. 5872, pp. 343–347, Jul. 1982.

[75] I. Boulalas, A. Zaravinos, I. Karyotis, D. Delakas, and D. A. Spandidos, "Activation of RAS Family Genes in Urothelial Carcinoma," *J. Urol.*, vol. 181, no. 5, pp. 2312–2319, 2009.

[76] C. P. N. Dinney, D. J. McConkey, R. E. Millikan, X. Wu, M. Bar-Eli, L. Adam, A. M. Kamat, A. O. Siefker-Radtke, T. Tuziak, A. L. Sabichi, H. B. Grossman, W. F. Benedict, and B. Czerniak,

"Focus on bladder cancer," *Cancer Cell*, vol. 6, no. 2, pp. 111–116, 2004.

[77] R. Shepherd, S. A. Forbes, D. Beare, S. Bamford, C. G. Cole, S. Ward, N. Bindal, P. Gunasekaran, M. Jia, C. Y. Kok, K. Leung, A. Menzies, A. P. Butler, J. W. Teague, P. J. Campbell, M. R. Stratton, and P. A. Futreal, "Data mining using the Catalogue of Somatic Mutations in Cancer BioMart," *Database*, vol. 2011, p. bar018-bar018, Jan. 2011.

[78] B. L. Cheng, R. M. Neumann, A. L. Weaver, B. E. Spotts, and D. G. Bostwick, "Bladder Carcinoma," vol. 17, no. 10, pp. 3182–3187, 1999.

[79] D. Vageli, H. Kiaris, D. Delakas, P. Anezinis, A. Cranidis, and D. A. Spandidos, "Transcriptional activation of H-ras, K-ras and N-ras proto-oncogenes in human bladder tumors," *Cancer Lett.*, vol. 107, no. 2, pp. 241–247, 1996.

[80] A. Membrino, S. Cogoi, E. B. Pedersen, and L. E. Xodo, "G4-DNA formation in the HRAS promoter and rational design of decoy oligonucleotides for cancer therapy," *PLoS One*, vol. 6, no. 9, 2011.

[81] S. Cogoi, S. Zorzet, A. E. Shchekotikhin, and L. E. Xodo, "Potent Apoptotic Response Induced by Chloroacetamidine Anthrathiophenediones in Bladder Cancer Cells," *J. Med. Chem.*, vol. 58, no. 14, pp. 5476–5485, 2015.

[82] D. P. Ryan, T. S. Hong, and N. Bardeesy, "Pancreatic Adenocarcinoma," pp. 1039–1049, 2014.

[83] R. Siegel, J. Ma, Z. Zou, and A. Jemal, "Cancer Statistics , 2014," vol. 64, no. 1, pp. 9–29, 2014.

[84] I. T. Konstantinidis, A. L. Warshaw, J. N. Allen, L. S. Blaszkowsky, C. F.-D. Castillo, V. Deshpande, T. S. Hong, E. L. Kwak, G. Y. Lauwers, D. P. Ryan, J. A. Wargo, K. D. Lillemoe, and C. R. Ferrone, "Pancreatic ductal adenocarcinoma: is there a survival difference for R1 resections versus locally advanced unresectable tumors? What is a 'true' R0 resection?," *Ann. Surg.*, vol. 257, no. 4, p. 731—736, Apr. 2013.

[85] V. T. H. B. M. Smit, A. J. M. Boot, A. M. M. Smits, G. J. Fleuren, C. J. Cornelisse, and J. L. Bos, "KRAS codon 12 mutations occur very frequently in pancreatic adenocarcinomas," *Nucleic Acids Res.*, vol. 16, no. 16, pp. 7773–7782, Aug. 1988.

[86] C. Almoguera, D. Shibata, K. Forrester, J. Martin, N. Arnheim, and M. Perucho, "Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes," *Cell*, vol. 53, no. 4, pp. 549–554, 1988.

[87] S. Jones, S. Jones, X. Zhang, D. W. Parsons, J. C. Lin, R. J. Leary, P. Angenendt, P. Mankoo, H. Carter, H. Kamiyama, A. Jimeno, S. Hong, B. Fu, M. Lin, E. S. Calhoun, M. Kamiyama, K. Walter, T. Nikolskaya, Y. Nikolsky, J. Hartigan, D. R. Smith, M. Hidalgo, S. D. Leach, A. P. Klein, E. M. Jaffee, M. Goggins, A. Maitra, C. Iacobuzio-donahue, J. R. Eshleman, S. E. Kern, R. H. Hruban, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V. E. Velculescu, and K. W. Kinzler, "Core Signaling Pathways in Human," vol. 1801, 2008.

[88] A. J. Aguirre, N. Bardeesy, M. Sinha, L. Lopez, D. A. Tuveson, J. Horner, M. S. Redston, and R. A. Depinho, "Activated Kras and Ink4a / Arf deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma," pp. 3112–3126, 2003.

[89] A. Singh, P. Greninger, D. Rhodes, L. Koopman, S. Violette, and N. Bardeesy, "A Gene Expression Signature Associated with '" K-Ras Addiction "' Reveals Regulators of EMT and Tumor Cell Survival," *Cancer Cell*, vol. 15, no. 6, pp. 489–500, 2009.

[90] M. A. Collins, F. Bednar, Y. Zhang, J. Brisset, S. Galbán, C. J. Galbán, S. Rakshit, K. S. Flannagan, N. V. Adsay, and M. Pasca, "Oncogenic Kras is required for both the initiation and maintenance of pancreatic cancer in mice," vol. 26, no. 16.

[91] H. Ying, A. C. Kimmelman, C. a. Lyssiotis, S. Hua, G. C. Chu, E. Fletcher-Sananikone, J. W. Locasale, J. Son, H. Zhang, J. L. Coloff, H. Yan, W. Wang, S. Chen, A. Viale, H. Zheng, J. H. Paik, C. Lim, A. R. Guimaraes, E. S. Martin, J. Chang, A. F. Hezel, S. R. Perry, J. Hu, B. Gan, Y. Xiao, J. M. Asara, R. Weissleder, Y. A. Wang, L. Chin, L. C. Cantley, and R. a. Depinho, "Oncogenic kras maintains pancreatic tumors through regulation of anabolic glucose metabolism," *Cell*, vol. 149, no. 3, pp. 656–670, 2012.

[92] G. M. DeNicola and D. A. Tuveson, "RAS in cellular transformation and senescence," *Eur. J. Cancer*, vol. 45, no. SUPPL. 1, pp. 211–216, Oct. 2009.

[93] C. Guerra, M. Collado, C. Navas, A. J. Schuhmacher, I. Hernández-Porras, M. Cañamero, M. Rodriguez-Justo, M. Serrano, and M. Barbacid, "Pancreatitis-Induced Inflammation Contributes to Pancreatic Cancer by Inhibiting Oncogene-Induced Senescence," *Cancer Cell*, vol. 19, no. 6, pp. 728–739, Oct. 2017.

[94] A. Singh, P. Greninger, D. Rhodes, L. Koopman, S. Violette, N. Bardeesy, and J. Settleman, "A Gene Expression Signature Associated with 'K-Ras Addiction' Reveals Regulators of EMT and Tumor Cell Survival," *Cancer Cell*, vol. 15, no. 6, pp. 489–500, Jun. 2009.

[95] L. Elghazi, A. J. Weiss, D. J. Barker, J. Callaghan, L. Staloch, E. P. Sandgren, M. Gannon, V. N. Adsay, and E. Bernal–Mizrachi, "Regulation of Pancreas Plasticity and Malignant Transformation by Akt Signaling," *Gastroenterology*, vol. 136, no. 3, p. 1091–1103.e8, Mar. 2009.

[96] S. Eser, N. Reiff, M. Messer, B. Seidler, K. Gottschalk, M. Dobler, M. Hieber, A. Arbeiter, S. Klein, B. Kong, C. W. Michalski, A. M. Schlitter, I. Esposito, A. J. Kind, L. Rad, A. E. Schnieke, M. Baccarini, D. R. Alessi, R. Rad, R. M. Schmid, G. Schneider, and D. Saur, "Selective Requirement of PI3K/PDK1 Signaling for Kras Oncogene-Driven Pancreatic Cell Plasticity and Cancer," *Cancer Cell*, vol. 23, no. 3, pp. 406–420, Mar. 2013.

[97] I. B. Weinstein, "Disorders in cell circuitry during multistage carcinogenesis: the role of homeostasis," *Carcinogenesis*, vol. 21, no. 5, pp. 857–864, 2000.

[98] I. B. Weinstein and A. Joe, "Oncogene addiction," *Cancer Res.*, vol. 68, no. 9, pp. 3077–3080, 2008.

[99] I. B. Weinstein, "Cancer. Addiction to oncogenes--the Achilles heal of cancer," *Science*, vol. 297, no. 5578, p. 63—64, 2002.

[100] D. W. Felsher and J. M. Bishop, "Reversible tumorigenesis by MYC in hematopoietic lineages," *Mol. Cell*, vol. 4, no. 2, pp. 199–207, 1999.

[101] M. Jain, C. Arvanitis, K. Chu, W. Dewey, E. Leonhardt, M. Trinh, C. D. Sundberg, J. M. Bishop, and D. W. Felsher, "Sustained Loss of a Neoplastic Phenotype by Brief Inactivation of MYC," *Science (80-. ).*, vol. 297, no. 5578, pp. 102–104, 2002.

[102] T. Holien, T. K. Våtsveen, H. Hella, A. Waage, A. Sundan, W. Dc, T. Holien, T. K. Våtsveen, H. Hella, A. Waage, and A. Sundan, "Addiction to c-MYC in multiple myeloma," vol. 120, no. 12, pp. 2450–2453, 2013.

[103] L. Chin, A. Tam, J. Pomerantz, M. Wong, J. Holash, N. Bardeesy, Q. Shen, R. O'Hagan, J. Pantginis, H. Zhou, J. W. Horner, C. Cordon-Cardo, G. D. Yancopoulos, and R. A. DePinho,

"Essential role for oncogenic Ras in tumour maintenance," *Nature*, vol. 400, no. 6743, p. 468—472, 1999.

[104] S. Shirasawa, M. Furuse, N. Yokoyama, and T. Sasazuki, "Altered Growth of Human Colon Cancer Cell Lines Disrupted at Activated Ki-ras," *Science (80-. ).*, vol. 260, no. 5104, pp. 85–88, 2009.

[105] M. A. Collins, F. Bednar, Y. Zhang, J. Brisset, S. Galbán, C. J. Galbán, S. Rakshit, K. S. Flannagan, N. V. Adsay, M. Pasca di Magliano, and M. Pasca, "Oncogenic Kras is required for both the initiation and maintenance of pancreatic cancer in mice," *J. Clin. Invest.*, vol. 122, no. 2, pp. 639–653, 2012.

[106] K. L. Bryant, J. D. Mancias, A. C. Kimmelman, and C. J. Der, "KRAS: Feeding pancreatic cancer proliferation," *Trends Biochem. Sci.*, vol. 39, no. 2, pp. 91–100, 2014.

[107] J. Son, C. A. Lyssiotis, H. Ying, X. Wang, S. Hua, M. Ligorio, R. M. Perera, C. R. Ferrone, E. Mullarky, N. Shyh-Chang, Y. Kang, J. B. Fleming, N. Bardeesy, J. M. Asara, M. C. Haigis, R. A. DePinho, L. C. Cantley, and A. C. Kimmelman, "Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway," *Nature*, vol. 496, no. 7443, pp. 101–105, 2013.

[108] H. Ledford, "The ras renaissance," *Nature*, vol. 520, pp. 278–280, 2015.

[109] V. González-Pérez, D. J. Reiner, J. K. Alan, C. Mitchell, L. J. Edwards, V. Khazak, C. J. Der, and A. D. Cox, "Genetic and functional characterization of putative Ras/Raf interaction inhibitors in C. elegans and mammalian cells.," *J. Mol. Signal.*, vol. 5, p. 2, 2010.

[110] J. Kato-Stankiewicz, I. Hakimi, G. Zhi, J. Zhang, I. Serebriiskii, L. Guo, H. Edamatsu, H. Koide, S. Menon, R. Eckl, S. Sakamuri, Y. Lu, Q.-Z. Chen, S. Agarwal, W. R. Baumbach, E. A. Golemis, F. Tamanoi, and V. Khazak, "Inhibitors of Ras/Raf-1 interaction identified by two-hybrid screening revert Ras-dependent transformation phenotypes in human cancer cells," *Proc. Natl. Acad. Sci.*, vol. 99, no. 22, pp. 14398–14403, 2002.

[111] D. Nayak, T. L. Roth, and D. B. Mcgavern, "An Orthosteric Inhibitor of the Ras-Sos Interaction," *Nat. Chem. Biol.*, vol. 7, no. 2, pp. 367–402, 2014.

[112] S. M. S. Norbert Berndt, Andrew D. Hamilton, "Targeting protein prenylation for cancer therapy," *Nat. Rev. Cancer*, vol. 11, no. 11, pp. 775–791, 2011.

[113] A. D. Cox, S. W. Fesik, A. C. Kimmelman, J. Luo, and C. J. Der, "Drugging the undruggable Ras: mission possible?," vol. 13, no. 11, pp. 828–851, 2015.

[114] A. J. Levine and A. M. Puzio-Kuter, "The Control of the Metabolic Switch in Cancers by Oncogenes and Tumor Suppressor Genes," *Science (80-. ).*, vol. 330, no. 6009, pp. 1340–1344, 2010.

[115] K. Shukla, D. V. Ferraris, A. G. Thomas, M. Stathis, B. Duvall, G. Delahanty, J. Alt, R. Rais, C. Rojas, P. Gao, Y. Xiang, C. V. Dang, B. S. Slusher, and T. Tsukamoto, "Design, synthesis, and pharmacological evaluation of bis-2-(5- phenylacetamido-1,2,4-thiadiazol-2-yl)ethyl sulfide 3 (BPTES) analogs as glutaminase inhibitors," *J. Med. Chem.*, vol. 55, no. 23, pp. 10551–10563, 2012.

[116] R. Perera, J. Son, C. A. Lyssiotis, H. Ying, X. Wang, S. Hua, M. Ligorio, R. M. Perera, C. R. Ferrone, E. Mullarky, and N. Shyh-Chang, "Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway." 2013.

[117] S. Cogoi, A. E. Shchekotikhin, and L. E. Xodo, "HRAS is silenced by two neighboring G-quadruplexes and activated by MAZ, a zinc-finger transcription factor with DNA unfolding property," *Nucleic Acids Res.*, vol. 42, no. 13, pp. 8379–8388, 2014.

[118] S. Cogoi, M. Paramasivam, A. Membrino, K. K. Yokoyama, and L. E. Xodo, "The KRAS promoter responds to Myc-associated zinc finger and poly(ADP-ribose) polymerase 1 proteins, which recognize a critical quadruplex-forming GA-element," *J. Biol. Chem.*, vol. 285, no. 29, pp. 22003–22016, 2010.

[119] M. Paramasivam, A. Membrino, S. Cogoi, H. Fukuda, H. Nakagama, and L. E. Xodo, "Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: Implications for transcription," *Nucleic Acids Res.*, vol. 37, no. 9, pp. 2841–2853, 2009.

[120] S. Cogoi, S. Zorzet, V. Rapozzi, I. Géci, E. B. Pedersen, and L. E. Xodo, "MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice," *Nucleic Acids Res.*, vol. 41, no. 7, pp. 4049–4064, 2013.

[121] S. Cogoi, U. Jakobsen, E. B. Pedersen, S. Vogel, and L. E. Xodo, "Lipid-modified G4-decoy oligonucleotide anchored to nanoparticles: delivery and bioactivity in pancreatic cancer cells," *Sci. Rep.*, vol. 6, no. November, p. 38468, 2016.

[122] E. Huntzinger and E. Izaurralde, "Gene silencing by microRNAs: contributions of translational repression and mRNA decay," *Nat. Rev. Genet.*, vol. 12, no. 2, p. 99—110, 2011.

[123] H. oki Iwakawa and Y. Tomari, "The Functions of MicroRNAs: mRNA Decay and Translational Repression," *Trends Cell Biol.*, vol. 25, no. 11, pp. 651–665, 2015.

[124] R. Garzon, G. A. Calin, and C. M. Croce, "MicroRNAs in Cancer," *Annu. Rev. Med.*, vol. 60, no. 1, pp. 167–179, Feb. 2009.

[125] F. J. Slack and J. B. Weidhaas, "MicroRNA in cancer prognosis," *N. Engl. J. Med.*, vol. 359, no. 25, p. 2720—2722, 2008.

[126] S. M. Johnson, H. Grosshans, J. Shingara, M. Byrom, R. Jarvis, A. Cheng, E. Labourier, K. L. Reinert, D. Brown, and F. J. Slack, "RAS is regulated by the let-7 microRNA family," *Cell*, vol. 120, no. 5, pp. 635–647, 2005.

[127] S. Yu, Z. Lu, C. Liu, Y. Meng, and Y. Ma, "miRNA-96 Suppresses KRAS and Functions as a Tumor Suppressor Gene in Pancreatic Cancer," pp. 6015–6025.

[128] M. Deng, H. Tang, Y. Zhou, M. Zhou, W. Xiong, Y. Zheng, Q. Ye, X. Zeng, Q. Liao, X. Guo, X. Li, J. Ma, and G. Li, "miR-216b suppresses tumor growth and invasion by targeting KRAS in nasopharyngeal carcinoma.," *J. Cell Sci.*, vol. 124, no. Pt 17, pp. 2997–3005, 2011.

[129] A. Osorio, J. Benitez, and B. Martinez-delgado, "Deregulated miRNAs in Hereditary Breast Cancer Revealed a Role for miR-30c in Regulating KRAS Oncogene," vol. 7, no. 6, 2012.

[130] M. Matsui and D. R. Corey, "Non-coding RNAs as drug targets," *Nat. Rev. Drug Discov.*, vol. 16, no. 3, pp. 167–179, 2016.

[131] H. Ling, M. Fabbri, and G. A. Calin, "MicroRNAs and other non-coding RNAs as targets for anticancer drug development," *Nat. Rev. Drug Discov.*, vol. 12, no. 11, pp. 847–865, 2013.

[132] N. Langkjær, A. Pasternak, and J. Wengel, "UNA (unlocked nucleic acid): A flexible RNA

mimic that allows engineering of nucleic acid duplex stability," *Bioorganic Med. Chem.*, vol. 17, no. 15, pp. 5420–5425, 2009.

[133] Q. Ge, A. Dallas, H. Ilves, J. Shorenstein, M. A. Behlke, and B. H. Johnston, "Effects of chemical modification on the potency, serum stability, and immunostimulatory properties of short shRNAs," *Rna*, vol. 16, no. 1, pp. 118–130, 2010.

[134] I. Bang, "Untersuchungen über die Guanylsäure," *Biochem. Z.*, vol. 26, pp. 293–311, 1910.

[135] I. Gellert, M. N. Lipsett, and D. R. Davies, "Helix formation by guanylic acid," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 48, pp. 2013–2018, 1962.

[136] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, and S. Neidle, "Quadruplex DNA: Sequence, topology and structure," *Nucleic Acids Res.*, vol. 34, no. 19, pp. 5402–5415, 2006.

[137] J. L. Huppert and S. Balasubramanian, "Prevalence of quadruplexes in the human genome," *Nucleic Acids Res.*, vol. 33, no. 9, pp. 2908–2916, 2005.

[138] A. Bedrat, L. Lacroix, and J. L. Mergny, "Re-evaluation of G-quadruplex propensity with G4Hunter," *Nucleic Acids Res.*, vol. 44, no. 4, pp. 1746–1759, 2016.

[139] N. Beaume, R. Pathak, V. K. Yadav, S. Kota, H. S. Misra, H. K. Gautam, and S. Chowdhury, "Genome-wide study predicts promoter-G4 DNA motifs regulate selective functions in bacteria: Radioresistance of D. radiodurans involves G4 DNA-mediated regulation," *Nucleic Acids Res.*, vol. 41, no. 1, pp. 76–89, 2013.

[140] M. Metifiot, S. Amrane, S. Litvak, and M.-L. Andreola, "G-quadruplexes in viruses: function and potential therapeutic applications," *Nucleic Acids Res.*, vol. 42, no. 20, pp. 12352–12366, 2014.

[141] C. Schaffitzel, I. Berger, J. Postberg, J. Hanes, H. J. Lipps, and A. Plückthun, "In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 15, pp. 8572–7, 2001.

[142] A. Henderson, Y. Wu, Y. C. Huang, E. A. Chavez, J. Platt, F. B. Johnson, R. M. Brosh, D. Sen, and P. M. Lansdorp, "Detection of G-quadruplex DNA in mammalian cells," *Nucleic Acids Res.*, vol. 42, no. 2, pp. 860–869, 2014.

[143] G. Biffi, D. Tannahill, J. McCafferty, and S. Balasubramanian, "Quantitative visualization of DNA G-quadruplex structures in human cells," *Nat. Chem.*, vol. 5, no. 3, pp. 182–186, 2013.

[144] H. G. Kazemier, K. Paeschke, and P. M. Lansdorp, "Guanine quadruplex monoclonal antibody 1H6 cross-reacts with restrained thymidine-rich single stranded DNA," *Nucleic Acids Res.*, vol. 45, no. 10, pp. 5913–5919, 2017.

[145] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, and S. Balasubramanian, "High-throughput sequencing of DNA G-quadruplex structures in the human genome," *Nat. Biotechnol.*, vol. 33, no. July, pp. 1–7, 2015.

[146] R. Hänsel-Hertsch, D. Beraldi, S. V Lensing, G. Marsico, K. Zyner, A. Parry, M. Di Antonio, J. Pike, H. Kimura, M. Narita, D. Tannahill, and S. Balasubramanian, "G-quadruplex structures mark human regulatory chromatin.," *Nat. Genet.*, vol. 48, no. 10, pp. 1267–72, 2016.

[147] D. Rhodes and H. J. Lipps, "Survey and summary G-quadruplexes and their regulatory roles in biology," *Nucleic Acids Res.*, vol. 43, no. 18, pp. 8627–8637, 2015.

[148] L. Hayflick, “The limited in vitro lifetime of human diploid cell strains,” *Exp. Cell Res.*, vol. 37, no. 3, pp. 614–636, 1965.

[149] C. W. Greider and E. H. Blackburn, “Identification of a specific telomere terminal transferase activity in tetrahymena extracts,” *Cell*, vol. 43, no. 2 PART 1, pp. 405–413, 1985.

[150] N. W. Kim, M. A. Piatyszek, K. R. Prowse, C. B. Harley, M. D. West, P. L. C. Ho, G. M. Coviello, W. E. Wright, S. L. Weinrich, and J. W. Shay, “Specific Association of Human Telomerase Activity with Immortal Cells and Cancer,” *Science (80-. ).*, vol. 266, no. 5193, pp. 2011–2015, 2016.

[151] R. K. Moyzis, J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff, and J. R. Wu, “A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 85, no. 18, pp. 6622–6, 1988.

[152] G. N. Parkinson, M. P. H. Lee, and S. Neidle, “Crystal structure of parallel quadruplexes from human telomeric DNA.,” *Nature*, vol. 417, no. 6891, pp. 876–80, 2002.

[153] Y. Wang and D. J. Patel, “Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex,” *Structure*, vol. 1, no. 4, pp. 263–282, 1993.

[154] H.-L. Bao, T. Ishizuka, T. Sakamoto, K. Fujimoto, T. Uechi, N. Kenmochi, and Y. Xu, “Characterization of human telomere RNA G-quadruplex structures in vitro and in living cells using 19F NMR spectroscopy,” *Nucleic Acids Res.*, vol. 45, no. 9, pp. 1–11, 2017.

[155] H. T. Le, W. L. Dean, R. Buscaglia, J. B. Chaires, and J. O. Trent, “An Investigation of G‐Quadruplex Structural Polymorphism in the Human Telomere Using a Combined Approach of Hydrodynamic Bead Modeling and Molecular Dynamics Simulation,” *J. Phys. Chem. B*, vol. 118, pp. 5390–5405, 2014.

[156] Anh Tuân Phan and Dinshaw J. Patel, “Two-Repeat Human Telomeric d(TAGGGTTAGGGT) Sequence Forms Interconverting Parallel and Antiparallel G-Quadruplexes in Solution: Distinct Topologies, Thermodynamic Properties, and Folding/Unfolding Kinetics Anh,” *J. Am. Chem. Soc.*, vol. 2, no. 74, pp. 15021–15027, 2015.

[157] A. T. Phan, K. N. Luu, and D. J. Patel, “Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K+ solution,” *Nucleic Acids Res.*, vol. 34, no. 19, pp. 5715–5719, 2006.

[158] D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle, and L. Hurley, “Inhibition of human telomerase by a G-Quadruplex-Interactive compound,” *J. Med. Chem.*, vol. 40, no. 14, pp. 2113–2116, 1997.

[159] A. M. Zahler, J. R. Williamson, T. R. Cech, and D. M. Prescott, “Inhibition of telomerase by G-quartet DNA structures,” *Nature*, vol. 350, no. 6320, pp. 718–720, 1991.

[160] J. F. Riou, L. Guittat, P. Mailliet, A. Laoui, E. Renou, O. Petitgenet, F. Mégnin-Chanet, C. Hélène, and J. L. Mergny, “Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 5, pp. 2672–2677, 2002.

[161] M. a Shammas, R. J. Shmookler Reis, C. Li, H. Koley, L. Hurley, K. C. Anderson, and N. C. Munshi, “Telomerase inhibition and cell growth arrest after telomestatin treatment in multiple myeloma.,” *Clin. Cancer Res.*, vol. 10, no. 2, pp. 770–776, 2004.

[162] R. A. Heald, C. Modi, J. C. Cookson, I. Hutchinson, C. A. Laughton, S. M. Gowan, L. R. Kelland, and M. F. G. Stevens, "Activity of Methylated Pentacyclic Acridinium Salts," pp. 590–597, 2002.

[163] S. K. Noureini, H. Esmaeili, F. Abachi, S. Khiali, B. Islam, M. Kuta, A. A. Saboury, M. Hoffmann, J. Sponer, G. Parkinson, and S. Haider, "Selectivity of major isoquinoline alkaloids from Chelidonium majus towards telomeric G-quadruplex: A study using a transition-FRET (t-FRET) assay," *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1861, no. 8, pp. 2020–2030, 2017.

[164] A. Kerkour, J. L. Mergny, and G. F. Salgado, "NMR based model of human telomeric repeat G-quadruplex in complex with 2,4,6-triarylpyridine family ligand," *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1861, no. 5, pp. 1293–1302, 2017.

[165] N. H. Campbell, G. N. Parkinson, A. P. Reszka, and S. Neidle, "Structural basis of DNA quadruplex recognition by an acridine drug," *J. Am. Chem. Soc.*, vol. 130, no. 21, pp. 6722–6724, 2008.

[166] A. M. Burger, F. Dai, C. M. Schultes, A. P. Reszka, M. J. Moore, J. A. Double, and S. Neidle, "The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function," *Cancer Res.*, vol. 65, no. 4, pp. 1489–1496, 2005.

[167] G. Zhou, X. Liu, Y. Li, S. Xu, C. Ma, X. Wu, Y. Cheng, Z. Yu, G. Zhao, and Y. Chen, "Telomere targeting with a novel G-quadruplex-interactive ligand BRACO-19 induces T-loop disassembly and telomerase displacement in human glioblastoma cells.," *Oncotarget*, vol. 7, no. 12, 2016.

[168] S. Neidle, "Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer," *FEBS J.*, vol. 277, no. 5, pp. 1118–1125, 2010.

[169] E. Salvati, C. Leonetti, A. Rizzo, M. Scarsella, M. Mottolese, R. Galati, I. Sperduti, M. F. G. Stevens, M. D'Incalci, M. Blasco, G. Chiorino, S. Bauwens, B. Horard, E. Gilson, A. Stoppacciaro, G. Zupi, and A. Biroccio, "Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect," *J. Clin. Invest.*, vol. 117, no. 11, pp. 3236–3247, 2007.

[170] S. Müller, D. A. Sanders, M. Di Antonio, S. Matsis, J.-F. Riou, R. Rodriguez, and S. Balasubramanian, "Pyridostatin analogues promote telomere dysfunction and long-term growth inhibition in human cancer cells," *Org. Biomol. Chem.*, vol. 10, no. 32, p. 6537, 2012.

[171] P. Bedinger, M. Munn, and B. M. Alberts, "Sequence-specific pausing during in vitro DNA replication on double-stranded DNA templates.," *J. Biol. Chem.*, vol. 264, no. 28, pp. 16880–16886, 1989.

[172] K. Usdin and K. J. Woodford, "CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro," *Nucleic Acids Res.*, vol. 23, no. 20, pp. 4202–4209, 1995.

[173] K. J. Woodford, R. M. Howell, and K. Usdin, "A novel K+-dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes," *J. Biol. Chem.*, vol. 269, no. 43, pp. 27029–27035, 1994.

[174] O. Mendoza, A. Bourdoncle, J. B. Boulé, R. M. Brosh, and J. L. Mergny, "G-quadruplexes and helicases," *Nucleic Acids Res.*, vol. 44, no. 5, pp. 1989–2006, 2016.

[175] K. L. Katrin Paeschke, Matthew L. Bochman , P. Daniela Garcia Petr Cejka and S. C. K. and V. A. Z. Friedman, "Pif1 family helicases suppress genome instability at G- quadruplex motifs

Katrin," vol. 21, no. 10, pp. 3240–3248, 2015.

[176] Y. Wu, K. Shin-ya, and R. M. Brosh, "FANCJ Helicase Defective in Fanconia Anemia and Breast Cancer Unwinds G-Quadruplex DNA To Defend Genomic Stability," *Mol. Cell. Biol.*, vol. 28, no. 12, pp. 4116–4128, 2008.

[177] C. M. Sanders, "Human Pifı helicase is a G-quadruplex DNA-binding protein with G-quadruplex DNA-unwinding activity," *Biochem J*, vol. 430, no. 1, pp. 119–128, 2010.

[178] K. Paeschke, J. A. Capra, and V. A. Zakian, "DNA Replication through G-Quadruplex Motifs Is Promoted by the Saccharomyces cerevisiae Pifı DNA Helicase," *Cell*, vol. 145, no. 5, pp. 678–691, 2011.

[179] T. B. C. London, L. J. Barber, G. Mosedale, G. P. Kelly, S. Balasubramanian, I. D. Hickson, S. J. Boulton, and K. Hiom, "FANCJ is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts," *J. Biol. Chem.*, vol. 283, no. 52, pp. 36132–36139, 2008.

[180] P. Castillo Bosch, S. Segura-Bayona, W. Koole, J. T. van Heteren, J. M. Dewar, M. Tijsterman, and P. Knipscheer, "FANCJ promotes DNA synthesis through G-quadruplex structures.," *EMBO J.*, vol. 33, no. 21, pp. 2521–33, 2014.

[181] J. L. Huppert and S. Balasubramanian, "G-quadruplexes in promoters throughout the human genome," *Nucleic Acids Res.*, vol. 35, no. 2, pp. 406–413, 2007.

[182] R. Perrone, E. Lavezzo, E. Riello, R. Manganelli, G. Palù, S. Toppo, R. Provvedi, and S. N. Richter, "Mapping and characterization of G-quadruplexes in Mycobacterium tuberculosis gene promoter regions," *Sci. Rep.*, vol. 7, no. 1, p. 5743, 2017.

[183] S. G. Hershman, Q. Chen, J. Y. Lee, M. L. Kozak, P. Yue, L. S. Wang, and F. B. Johnson, "Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in Saccharomyces cerevisiae," *Nucleic Acids Res.*, vol. 36, no. 1, pp. 144–156, 2008.

[184] J. Eddy, A. C. Vallur, S. Varma, H. Liu, W. C. Reinhold, Y. Pommier, and N. Maizels, "G4 motifs correlate with promoter-proximal transcriptional pausing in human genes," *Nucleic Acids Res.*, vol. 39, no. 12, pp. 4975–4983, 2011.

[185] M. L. Bochman, K. Paeschke, and V. A. Zakian, "DNA secondary structures: stability and function of G-quadruplex structures," *Nat Rev Genet*, vol. 13, no. 11, pp. 770–780, Nov. 2012.

[186] J. Eddy and N. Maizels, "Gene function correlates with potential for G4 DNA formation in the human genome," *Nucleic Acids Res.*, vol. 34, no. 14, pp. 3887–3896, 2006.

[187] T. A. Brooks and L. Hurley, "Targeting MYC Expression through G-Quadruplexes," *Genes Cancer*, vol. 1, no. 6, pp. 641–649, 2010.

[188] K. B. Marcu, S. A. Bossone, and A. J. Patel, "myc Function and Regulation," *Annu. Rev. Biochem.*, vol. 61, no. 1, pp. 809–858, Jun. 1992.

[189] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, and L. Hurley, "Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 18, pp. 11593–8, 2002.

[190] A. T. Phan, Y. S. Modi, and D. J. Patel, "Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter," *J. Am. Chem. Soc.*, vol. 126, no. 28, pp. 8710–8716, 2004.

[191] S. Balasubramanian and L. Hurley, "Targeting G-quadruplexes in gene promoters : a novel

anticancer," *Nat. Rev. Drug Discov.*, vol. 10, no. 4, pp. 261–275, 2011.

[192] F. Yamamoto and M. Perucho, "Characterization of the human cK-ras gene promoter.," *Oncogene Res.*, vol. 3, no. 2, pp. 125–130, 1988.

[193] S. Cogoi, M. Paramasivam, B. Spolaore, and L. E. Xodo, "Structural polymorphism within a regulatory element of the human KRAS promoter: Formation of G4-DNA recognized by nuclear proteins," *Nucleic Acids Res.*, vol. 36, no. 11, pp. 3765–3780, 2008.

[194] R. K. Morgan, H. Batra, V. C. Gaerig, J. Hockings, and T. A. Brooks, "Identification and characterization of a new G-quadruplex forming region within the kRAS promoter as a transcriptional regulator," *Biochim. Biophys. Acta - Gene Regul. Mech.*, vol. 1859, no. 2, pp. 235–245, 2016.

[195] J. Jean-Philippe, S. Paz, and M. Caputi, "hnRNP A1: The Swiss Army Knife of gene expression," *Int. J. Mol. Sci.*, vol. 14, no. 9, pp. 18999–19024, 2013.

[196] A. C. Krüger, M. K. Raarup, M. M. Nielsen, M. Kristensen, F. Besenbacher, J. Kjems, and V. Birkedal, "Interaction of hnRNP A1 with telomere DNA G-quadruplex structures studied at the single molecule level," *Eur. Biophys. J.*, vol. 39, no. 9, pp. 1343–1350, 2010.

[197] H. Fukuda, M. Katahira, N. Tsuchiya, Y. Enokizono, T. Sugimura, M. Nagao, and H. Nakagama, "Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 20, pp. 12685–90, 2002.

[198] P.-C. Chu, M.-C. Yang, S. K. Kulp, S. B. Salunke, L. E. Himmel, C.-S. Fang, A. M. Jadhav, Y.-S. Shan, C.-T. Lee, M.-D. Lai, L. A. Shirley, T. Bekaii-Saab, and C.-S. Chen, "Regulation of oncogenic KRAS signaling via a novel KRAS-integrin-linked kinase-hnRNPA1 regulatory loop in human pancreatic cancer cells," *Oncogene*, vol. 35, no. 30. Macmillan Publishers Limited, pp. 3897–3908, 28-Jul-2016.

[199] M. Paramasivam, S. Cogoi, and L. E. Xodo, "Primer extension reactions as a tool to uncover folding motifs within complex G-rich sequences: analysis of the human KRAS NHE.," *Chem. Commun. (Camb).*, vol. 47, no. 17, pp. 4965–7, 2011.

[200] A. Kerkour, J. Marquevielle, S. Ivashchenko, L. A. Yatsunyk, J. L. Mergny, and G. F. Salgado, "High-resolution three-dimensional NMR structure of the KRAS proto-oncogene promoter reveals key features of a G-quadruplex involved in transcriptional regulation," *J. Biol. Chem.*, vol. 292, no. 19, pp. 8082–8091, 2017.

[201] C. E. Kaiser, N. A. Van Ert, P. Agrawal, R. Chawla, D. Yang, and J. L. Huppert, "Insight into the Complexity of the i-Motif and G-Quadruplex DNA Structures Formed in the KRAS Promoter and Subsequent Drug-Induced Gene Repression," *J. Am. Chem. Soc.*, vol. 139, no. 25, pp. 8522–8536, 2017.

[202] Y. Yarden, W. J. Kuang, T. Yang-Feng, L. Coussens, S. Munemitsu, T. J. Dull, E. Chen, J. Schlessinger, U. Francke, and A. Ullrich, "Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an unidentified ligand.," *EMBO J.*, vol. 6, no. 11, pp. 3341–51, 1987.

[203] D. E. Williams, J. Eisenman, A. Baird, C. Rauch, K. Van Ness, C. J. March, L. S. Park, U. Martin, D. Y. Mochizukl, H. S. Boswell, G. S. Burgess, D. Cosman, and S. D. Lyman, "Identification of a ligand for the c-kit proto-oncogene," *Cell*, vol. 63, no. 1, pp. 167–174, 1990.

[204] S. Hirota, "Gain-of-Function Mutations of c-kit in Human Gastrointestinal Stromal Tumors," *Science (80-. ).*, vol. 279, no. 5350, pp. 577–580, 1998.

[205] C. M. V de Silva and R. Reid, "Gastrointestinal stromal tumors (GIST): C-kit mutations, CD117 expression, differential diagnosis and targeted cancer therapy with Imatinib.," *Pathol. Oncol. Res.*, vol. 9, no. 1, pp. 13–9, 2003.

[206] M. C. Heinrich, D. J. Griffith, B. J. Druker, C. L. Wait, K. a Ott, and a J. Zigler, "Inhibition of c-kit receptor tyrosine kinase activity by STI 571, a selective tyrosine kinase inhibitor.," *Blood*, vol. 96, no. 3, pp. 925–32, 2000.

[207] G. D. Demetri, A. T. van Oosterom, C. R. Garrett, M. E. Blackstein, M. H. Shah, J. Verweij, G. McArthur, I. R. Judson, M. C. Heinrich, J. A. Morgan, J. Desai, C. D. Fletcher, S. George, C. L. Bello, X. Huang, C. M. Baum, and P. G. Casali, "Efficacy and safety of sunitinib in patients with advanced gastrointestinal stromal tumour after failure of imatinib: a randomised controlled trial," *Lancet*, vol. 368, no. 9544, pp. 1329–1338, 2006.

[208] M. E. Gorre, "Clinical Resistance to STI-571 Cancer Therapy Caused by BCR-ABL Gene Mutation or Amplification," *Science (80-. ).*, vol. 293, no. 5531, pp. 876–880, 2001.

[209] D. Wei, J. Husby, and S. Neidle, "Flexibility and structural conservation in a c-KIT G-quadruplex," *Nucleic Acids Res.*, vol. 43, no. 1, pp. 629–644, 2015.

[210] S. T. D. Hsu, P. Varnai, A. Bugaut, A. P. Reszka, S. Neidle, and S. Balasubramanian, "A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics," *J. Am. Chem. Soc.*, vol. 131, no. 37, pp. 13399–13409, 2009.

[211] D. Wei, G. N. Parkinson, A. P. Reszka, and S. Neidle, "Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation," *Nucleic Acids Res.*, vol. 40, no. 10, pp. 4691–4700, 2012.

[212] V. Kuryavyi, A. T. Phan, and D. J. Patel, "Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter," *Nucleic Acids Res.*, vol. 38, no. 19, pp. 6757–6773, 2010.

[213] E. A. Raiber, R. Kranaster, E. Lam, M. Nikan, and S. Balasubramanian, "A non-canonical DNA structure is a binding motif for the transcription factor SP1 in vitro," *Nucleic Acids Res.*, vol. 40, no. 4, pp. 1499–1508, 2012.

[214] K. Yamamoto, A. Tojo, N. Aoki, and M. Shibuya, "Characterization of the promoter region of the human c-kit proto-oncogene.," *Japanese journal of cancer research : Gann*, vol. 84, no. 11. pp. 1136–44, 1993.

[215] R. Rigo and C. Sissi, "Characterization of G4–G4 Crosstalk in the *c-KIT* Promoter Region," *Biochemistry*, p. acs.biochem.7b00660, 2017.

[216] M. Bejugam, S. Sewitz, P. S. Shirude, R. Rodriguez, R. Shahid, and S. Balasubramanian, "Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: Small molecule regulation of c-kit oncogene expression," *J. Am. Chem. Soc.*, vol. 129, no. 43, pp. 12926–12927, 2007.

[217] K. I. E. McLuckie, Z. A. E. Waller, D. A. Sanders, D. Alves, R. Rodriguez, J. Dash, G. J. McKenzie, A. R. Venkitaraman, and S. Balasubramanian, "G-Quadruplex-Binding benzo[a]phenoxazines down-regulate c-KIT expression in human gastric carcinoma cells," *J. Am. Chem. Soc.*, vol. 133, no. 8, pp. 2658–2663, 2011.

[218] E. Zorzan, S. Da Ros, C. Musetti, L. Z. Shahidian, N. F. R. Coelho, F. Bonsembiante, S. Létard, M. E. Gelain, M. Palumbo, P. Dubreuil, M. Giantin, C. Sissi, and M. Dacasto, "Screening of

candidate G-quadruplex ligands for the human c-KIT promotorial region and their effects in multiple in-vitro models.," *Oncotarget*, vol. 7, no. 16, pp. 21658–75, 2016.

[219] M. M. Fay, S. M. Lyons, and P. Ivanov, "RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms," *J. Mol. Biol.*, vol. 429, no. 14, pp. 2127–2147, 2017.

[220] H. Guiset Miserachs, D. Donghi, R. Börner, S. Johannsen, and R. K. O. Sigel, "Distinct differences in metal ion specificity of RNA and DNA G-quadruplexes," *J. Biol. Inorg. Chem.*, vol. 21, no. 8, pp. 975–986, 2016.

[221] D. J. Patel, A. T. Phan, and V. Kuryavyi, "Human telomere, oncogenic promoter and 5′-UTR G-quadruplexes: Diverse higher order DNA and RNA targets for cancer therapeutics," *Nucleic Acids Res.*, vol. 35, no. 22, pp. 7429–7455, 2007.

[222] J. L. Huppert, A. Bugaut, S. Kumari, and S. Balasubramanian, "G-quadruplexes: The beginning and end of UTRs," *Nucleic Acids Res.*, vol. 36, no. 19, pp. 6260–6268, 2008.

[223] C. K. Kwok, A. B. Sahakyan, and S. Balasubramanian, "Structural Analysis using SHALiPE to Reveal RNA G-Quadruplex Formation in Human Precursor MicroRNA," *Angew. Chemie - Int. Ed.*, vol. 55, no. 31, pp. 8958–8961, 2016.

[224] M. J. Morris, Y. Negishi, C. Pazsint, J. D. Schonhoft, and S. Basu, "An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES," *J. Am. Chem. Soc.*, vol. 132, no. 50, pp. 17831–17839, 2010.

[225] C. K. Kwok, G. Marsico, A. B. Sahakyan, V. S. Chambers, and S. Balasubramanian, "rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome," *Nat. Methods*, vol. 13, no. 10, pp. 841–844, 2016.

[226] J. U. Guo and D. P. Bartel, "RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria," *Science (80-. ).*, vol. 353, no. 6306, p. aaf5371-aaf5371, 2016.

[227] D. Boehringer and N. Ban, "Trapping the Ribosome to Control Gene Expression," *Cell*, vol. 130, no. 6, pp. 983–985, 2007.

[228] J. R. Babendure, J. L. Babendure, J. Ding, and R. Y. Tsien, "Control of mammalian translation by mRNA structure near caps Control of mammalian translation by mRNA structure near caps," *Rna*, no. Kozak 1994, pp. 851–861, 2006.

[229] J. Mandel, B. Ehresmann, and C. Ehresmann, "The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif speci ® cally to its mRNA via a purine quartet motif," *EMBO J.*, vol. 4813, no. 17, pp. 4803–4813, 2001.

[230] S. Kumari, A. Bugaut, J. L. Huppert, and S. Balasubramanian, "An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation," *Nat. Chem. Biol.*, vol. 3, no. 4, pp. 218–221, 2007.

[231] S. Kumari, A. Bugaut, and S. Balasubramanian, "Position and Stability Are Determining Factors for Translation Repression by an RNA G-Quadruplex-Forming Sequence within the 5 ′ UTR of the NRAS," pp. 12664–12669, 2008.

[232] A. Bugaut and S. Balasubramanian, "5-UTR RNA G-quadruplexes: Translation regulation and targeting," *Nucleic Acids Res.*, vol. 40, no. 11, pp. 4727–4741, 2012.

[233] S. Kumari, A. Bugaut, J. L. Huppert, and S. Balasubramanian, RNA G-quadruplex in the 5 ′

UTR of the NRAS proto- oncogene modulates translation," vol. 3, no. 4, pp. 218–221, 2008.

[234] R. Shahid, A. Bugaut, and S. Balasubramanian, "The BCL-2 5- untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression," *Biochemistry*, vol. 49, no. 38, pp. 8300–8306, 2010.

[235] A. Arora and B. Suess, "An RNA G-quadruplex in the 3' UTR of the proto-oncogene PIM1 represses translation," *RNA Biol.*, vol. 8, no. February 2015, pp. 802–805, 2011.

[236] E. Crenshaw, B. P. Leung, C. K. Kwok, M. Sharoni, K. Olson, N. P. Sebastian, S. Ansaloni, R. Schweitzer-Stenner, M. R. Akins, P. C. Bevilacqua, and A. J. Saunders, "Amyloid precursor protein translation is regulated by a 3'UTR guanine quadruplex," *PLoS One*, vol. 10, no. 11, pp. 1–18, 2015.

[237] A. Arora, M. Dutkiewicz, V. Scaria, M. Hariharan, S. Maiti, and J. Kurreck, "Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif.," *RNA*, vol. 14, no. 7, pp. 1290–1296, 2008.

[238] R. Kostadinov, N. Malhotra, M. Viotti, R. Shine, L. D'Antonio, and P. Bagga, "GRSDB: a database of quadruplex forming G-rich sequences in alternatively processed mammalian pre-mRNA sequences.," *Nucleic Acids Res.*, vol. 34, no. Database issue, pp. D119–D124, 2006.

[239] O. Kikin, Z. Zappala, L. D'Antonio, and P. S. Bagga, "GRSDB2 and GRS_UTRdb: Databases of quadruplex forming G-rich sequences in pre-mRNAs and mRNAs," *Nucleic Acids Res.*, vol. 36, no. SUPPL. 1, pp. 141–148, 2008.

[240] M. C. Didiot, Z. Tian, C. Schaeffer, M. Subramanian, J. L. Mandel, and H. Moine, "The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer," *Nucleic Acids Res.*, vol. 36, no. 15, pp. 4902–4912, 2008.

[241] D. Gomez, T. Lamarteleur, L. Lacroix, P. Mailliet, J. L. Mergny, and J. F. Riou, "Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing," *Nucleic Acids Res.*, vol. 32, no. 1, pp. 371–379, 2004.

[242] A. K. Byrd and K. D. Raney, "A parallel quadruplex DNA is bound tightly but unfolded slowly by Pif1 helicase," *J. Biol. Chem.*, vol. 290, no. 10, pp. 6482–6494, 2015.

[243] L. Federici, A. Arcovito, G. L. Scaglione, F. Scaloni, C. Lo Sterzo, A. Di Matteo, B. Falini, B. Giardina, and M. Brunori, "Nucleophosmin C-terminal leukemia-associated domain interacts with G-rich quadruplex forming DNA," *J. Biol. Chem.*, vol. 285, no. 48, pp. 37138–37149, 2010.

[244] M. A. Islam, S. D. Thomas, V. V. Murty, K. J. Sedoris, and D. M. Miller, "c-Myc quadruplex-forming sequence Pu-27 induces extensive damage in both telomeric and nontelomeric regions of DNA," *J. Biol. Chem.*, vol. 289, no. 12, pp. 8521–8531, 2014.

[245] V. González and L. H. Hurley, "The C-terminus of nucleolin promotes the formation of the c-MYC G-quadruplex and inhibits c-MYC promoter activity," *Biochemistry*, vol. 49, no. 45, pp. 9706–9714, 2010.

[246] S. L. Cree, R. Fredericks, A. Miller, F. G. Pearce, V. Filichev, C. Fee, and M. A. Kennedy, "DNA G-quadruplexes show strong interaction with DNA methyltransferases in vitro," *FEBS Lett.*, vol. 590, pp. 2870–2883, 2016.

[247] R. K. Thakur, P. Kumar, K. Halder, A. Verma, A. Kar, J. L. Parent, R. Basundra, A. Kumar, and S. Chowdhury, "Metastases suppressor NM23-H2 interaction with G-quadruplex DNA within

c-MYC promoter nuclease hypersensitive element induces c-MYC expression," *Nucleic Acids Res.*, vol. 37, no. 1, pp. 172–183, 2009.

[248] B. Guyen, C. M. Schultes, P. Hazel, J. Mann, and S. Neidle, "Synthesis and evaluation of analogues of 10H-indolo[3,2-b]quinoline as G-quadruplex stabilising ligands and potential inhibitors of the enzyme telomerase.," *Org. Biomol. Chem.*, vol. 2, pp. 981–988, 2004.

[249] V. Caprio, B. Guyen, Y. Opoku-Boahen, J. Mann, S. M. Gowan, L. M. Kelland, M. a Read, and S. Neidle, "A Novel Inhibitor of Human Telomerase Derived from 10H-Indolo [3, 2- b] quinoline," *Bioorg. Med. Chem. Lett.*, vol. 10, pp. 2063–2066, 2000.

[250] D. Monchaud and M.-P. Teulade-Fichou, "A hitchhiker's guide to G-quadruplex ligands," *Org. Biomol. Chem.*, vol. 6, no. 4, pp. 627–636, 2008.

[251] L. Martino, B. Pagano, I. Fotticchia, S. Neidle, and C. Giancola, "Shedding light on the interaction between TMPyP4 and human telomeric quadruplexes," *J. Phys. Chem. B*, vol. 113, no. 44, pp. 14779–14786, 2009.

[252] J. Seenisamy, S. Bashyam, V. Gokhale, H. Vankayalapati, D. Sun, A. Siddiqui-Jain, N. Streiner, K. Shin-ya, E. White, W. D. Wilson, and L. H. Hurley, "Design and synthesis of an expanded porphyrin that has selectivity for the c-MYC G-quadruplex structure," *J. Am. Chem. Soc.*, vol. 127, no. 9, pp. 2944–2959, 2005.

[253] W. M. Clark, W. Brooks, A. Mackey, M. D. Hill, P. P. Leimgruber, A. J. Sheffet, D. Ph, V. J. Howard, D. Ph, W. S. Moore, J. H. Voeks, D. Ph, L. N. Hopkins, D. E. Cutlip, D. J. Cohen, J. J. Popma, R. D. Ferguson, S. N. Cohen, J. L. Blackshear, F. L. Silver, J. P. Mohr, B. K. Lal, J. F. Meschia, and C. Investigators, "New England Journal," *N. Engl. J. Med.*, vol. 351, no. 21, pp. 2159–2169, 2004.

[254] D. Drygin, A. Siddiqui-Jain, S. O'Brien, M. Schwaebe, A. Lin, J. Bliesath, C. B. Ho, C. Proffitt, K. Trent, J. P. Whitten, J. K. C. Lim, D. Von Hoff, K. Anderes, and W. G. Rice, "Anticancer activity of CX-3543: A direct inhibitor of rRNA biogenesis," *Cancer Res.*, vol. 69, no. 19, pp. 7653–7661, 2009.

[255] H. Xu, M. Di Antonio, S. McKinney, V. Mathew, B. Ho, N. J. O'Neil, N. Dos Santos, J. Silvester, V. Wei, J. Garcia, F. Kabeer, D. Lai, P. Soriano, J. Banáth, D. S. Chiu, D. Yap, D. D. Le, F. B. Ye, A. Zhang, K. Thu, J. Soong, S. Lin, A. H. C. Tsai, T. Osako, T. Algara, D. N. Saunders, J. Wong, J. Xian, M. B. Bally, J. D. Brenton, G. W. Brown, S. P. Shah, D. Cescon, T. W. Mak, C. Caldas, P. C. Stirling, P. Hieter, S. Balasubramanian, and S. Aparicio, "CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours," *Nat. Commun.*, vol. 8, no. 205, p. 14432, 2017.

[256] K. Gehring, J.-L. Leroy, and M. Gueron, "A tetrameric DNA structure with protonated cytosine-cytosine base pairs," *Nature*, vol. 363, no. 6429, pp. 561–565, Jun. 1993.

[257] J. Laskar, F. Joutel, and P. Robutel, "© 19 9 3 Nature Publishing Group," *Nature*, vol. 361, 1993.

[258] D. Liu and S. Balasubramanian, "A Proton-Fuelled DNA Nanomachine," *Angew. Chemie - Int. Ed.*, vol. 42, no. 46, pp. 5734–5736, 2003.

[259] J. Sharma, R. Chhabra, H. Yan, and Y. Liu, "pH-driven conformational switch of 'i-motif' DNA for the reversible assembly of gold nanoparticles," *Chem. Commun.*, no. 5, pp. 477–479, 2007.

[260] W. Wang, H. Liu, D. Liu, Y. Xu, Y. Yang, and D. Zhou, "Use of the interparticle i-motif for the controlled assembly of gold nanoparticles," *Langmuir*, vol. 23, no. 24, pp. 11956–11959, 2007.

[261] S. Modi, M. G.Swetha, D. Goswami, G. D. Gupta, S. Mayor, and Y. Krishnan, "A DNA nanomachine that maps spatial and temporal pH changes inside living cells," *Nat Nano*, vol. 4, no. 5, pp. 325–330, May 2009.

[262] S. Surana, J. M. Bhat, S. P. Koushika, and Y. Krishnan, "An autonomous DNA nanomachine maps spatiotemporal pH changes in a multicellular living organism," *Nat. Commun.*, vol. 2, no. May, p. 340, 2011.

[263] T. Li, D. Ackermann, A. M. Hall, and M. Famulok, "Input-dependent induction of oligonucleotide structural motifs for performing molecular logic," *J. Am. Chem. Soc.*, vol. 134, no. 7, pp. 3508–3516, 2012.

[264] J. Elbaz, Z. G. Wang, R. Orbach, and L. Willner, "pH-stimulated concurrent mechanical activation of two DNA 'tweezers'. A 'SET-RESET' logic gate system," *Nano Lett.*, vol. 9, no. 12, pp. 4510–4514, 2009.

[265] R. E. Marsh, R. Bierstedt, and E. L. Eichhorn, "The crystal structure of cytosine-5-acetic acid," *Acta Crystallogr.*, vol. 15, no. 4, pp. 310–316, 1962.

[266] S. Ahmed and E. Henderson, "Formation of novel hairpin structures by telemeric C-strand oligonucleotides," *Nucleic Acids Res.*, vol. 20, no. 3, pp. 507–511, 1992.

[267] A. T. Phan and J.-L. Leroy, "Intramolecular i-Motif Structures of Telomeric DNA," *J. Biomol. Struct. Dyn.*, vol. 17, no. sup1, pp. 245–251, Jan. 2000.

[268] L. Cai, L. Chen, S. Raghavan, R. Ratliff, R. Moyzis, and A. Rich, "Intercalated cytosine motif and novel adenine clusters in the crystal structure of the Tetrahymena telomere," *Nucleic Acids Res.*, vol. 26, no. 20, pp. 4696–4705, 1998.

[269] J. Weil, T. Min, C. Yang, S. Wang, C. Sutherland, N. Sinha, and C. Kang, "Stabilization of the {\it i}-motif by intramolecular adenine{--}adenine{--}thymine base triple in the {\-}structure of d(ACCCT)," *Acta Crystallogr. Sect. D*, vol. 55, no. 2, pp. 422–429, Feb. 1999.

[270] N. Khan, A. Aviñó, R. Tauler, C. González, R. Eritja, and R. Gargallo, "Solution equilibria of the i-motif-forming region upstream of the B-cell lymphoma-2 P1 promoter," *Biochimie*, vol. 89, no. 12, pp. 1562–1572, 2007.

[271] T. A. Brooks, S. Kendrick, and L. Hurley, "Making sense of G-quadruplex and i-motif functions in oncogene promoters," *FEBS J.*, vol. 277, no. 17, pp. 3459–3469, 2010.

[272] L. Lacroix, J. L. Mergny, J. L. Leroy, and C. Hélène, "Inability of RNA to form the i-motif: Implications for triplex formation," *Biochemistry*, vol. 35, no. 26, pp. 8715–8722, 1996.

[273] S. Kendrick and L. H. Hurley, "The role of G-quadruplex/i-motif secondary structures as cis-acting regulatory elements.," *Pure Appl. Chem.*, vol. 82, no. 8, pp. 1609–1621, 2010.

[274] D. Sun and L. H. Hurley, "The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: Implications for drug targeting and control of gene expression," *J. Med. Chem.*, vol. 52, no. 9, pp. 2863–2874, 2009.

[275] D. Hanahan and R. a. Weinberg, "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

[276] M. Vander Heiden, L. Cantley, and C. Thompson, "Understanding the Warburg effect: The metabolic Requiremetns of cell proliferation," *Science (80-. ).*, vol. 324, no. 5930, pp. 1029–1033,

2009.

[277] E. B. Affar, R. G. Shah, A.-K. Dallaire, V. Castonguay, and G. M. Shah, "Role of poly(ADP-ribose) polymerase in rapid intracellular acidification induced by alkylating DNA damage.," *Proc Natl Acad Sci U S A*, vol. 99, no. 1, pp. 245–250, 2002.

[278] A. Rajendran, S. Nakano, and N. Sugimoto, "Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH," *Chem. Commun.*, vol. 46, no. 8, p. 1299, 2010.

[279] J. Cui, P. Waltman, V. H. Le, and E. A. Lewis, "The effect of molecular crowding on the stability of human c-MYC promoter sequence I-motif at neutral pH," *Molecules*, vol. 18, no. 10, pp. 12751–12767, 2013.

[280] J. J. Wenzel, H. Rossmann, C. Fottner, S. Neuwirth, C. Neukirch, P. Lohse, J. K. Bickmann, T. Minnemann, T. J. Musholt, B. Schneider-Rätzke, M. M. Weber, and K. J. Lackner, "Identification and prevention of genotyping errors caused by G-quadruplex- and i-motif-like sequences," *Clin. Chem.*, vol. 55, no. 7, pp. 1361–1371, 2009.

[281] G. Manzini, N. Yathindra, and L. E. Xodo, "Evidence for intramolecularly folded I-DNA structures in biologically relevant CCC-repeat sequences," *Nucleic Acids Res.*, vol. 22, no. 22, pp. 4634–4640, 1994.

[282] T. Simonsson, M. Pribylova, and M. Vorlickova, "A Nuclease Hypersensitive Element in the Human c-myc Promoter Adopts Several Distinct i-Tetraplex Structures," *Biochem. Biophys. Res. Commun.*, vol. 278, no. 1, pp. 158–166, 2000.

[283] S. Kendrick, Y. Akiyama, S. M. Hecht, and L. H. Hurley, "The i-motif in the bcl-2 P1 promoter forms an unexpectedly stable structure with a unique 8:5:7 loop folding pattern," *J. Am. Chem. Soc.*, vol. 131, no. 48, pp. 17667–17676, 2009.

[284] K. Guo, V. Gokhale, L. H. Hurley, and D. Sun, "Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene," *Nucleic Acids Res.*, vol. 36, no. 14, pp. 4598–4608, 2008.

[285] K. Guo, A. Pourpak, K. Beetz-Rogers, V. Gokhale, D. Sun, and L. H. Hurley, "Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene," *J. Am. Chem. Soc.*, vol. 129, no. 33, pp. 10220–10228, 2007.

[286] Y. Xu and H. Sugiyama, "Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb)," *Nucleic Acids Res.*, vol. 34, no. 3, pp. 949–954, 2006.

[287] L. Lacroix, H. Liénard, E. Labourier, M. Djavaheri-Mergny, J. Lacoste, H. Leffers, J. Tazi, C. Hélène, and J. L. Mergny, "Identification of two human nuclear proteins that recognise the cytosine-rich strand of human telomeres in vitro.," *Nucleic Acids Res.*, vol. 28, no. 7, pp. 1564–1575, 2000.

[288] H. Kang, S. Kendrick, S. M. Hecht, and L. H. Hurley, "Is a Molecular Switch for Control of Gene Expression That Can Be," 2014.

[289] V. González and L. H. Hurley, "The c-MYC NHE III1: Function and Regulation," *Annu. Rev. Pharmacol. Toxicol.*, vol. 50, no. 1, pp. 111–129, Jan. 2010.

[290] P. H. Backe, A. C. Messias, R. B. G. Ravelli, M. Sattler, and S. Cusack, "X-ray crystallographic and NMR studies of the third KH domain of hnRNP K in complex with single-stranded nucleic

acids," *Structure*, vol. 13, no. 7, pp. 1055–10607, 2005.

[291] J. F. Cornuel, A. Moraillon, and M. Guéron, "Participation of yeast inosine 5′-monophosphate dehydrogenase in an in vitro complex with a fragment of the C-rich telomeric strand," *Biochimie*, vol. 84, no. 4, pp. 279–289, 2002.

[292] G. Sarig, P. Weisman-Shomer, R. Erlitzki, and M. Fry, "Purification and characterization of qTBP42, a new single-stranded and quadruplex telomeric DNA-binding protein from rat hepatocytes," *J. Biol. Chem.*, vol. 272, no. 7, pp. 4474–4482, 1997.

[293] a Cortés and F. Azorín, "DDP1, a heterochromatin-associated multi-KH-domain protein of Drosophila melanogaster, interacts specifically with centromeric satellite DNA sequences.," *Mol. Cell. Biol.*, vol. 20, no. 11, pp. 3860–3869, 2000.